The shark strikes twice:

Generation of Mono- and Bi-specific High-Affinity vNAR Antibody Domains *via* Step-Wise Affinity Maturation

TECHNISCHE UNIVERSITÄT DARMSTADT

Vom Fachbereich Chemie
der Technischen Universität Darmstadt

zur Erlangung des akademischen Grades eines
Doctor rerum naturalium (Dr. rer. nat)

genehmigte
Dissertation

vorgelegt von

M.Sc. Stefan Zielonka
aus Viernheim

Referent:	Prof. Dr. Harald Kolmar
Korreferenten:	Prof. Dr. Stefan Dübel, Prof. Dr. Florian Rüker
Tag der Einreichung:	09. Februar 2015
Tag der mündlichen Prüfung:	31. März 2015

Darmstadt 2015

D17

Die vorliegende Arbeit wurde unter der Leitung von Herrn Prof. Dr. Harald Kolmar am Clemens-Schöpf-Institut für Organische Chemie und Biochemie der Technischen Universität Darmstadt von Januar 2012 bis Januar 2015 angefertigt.

Publications derived from this work

Parts of this work have been published or are currently under review in:

Zielonka S, Empting M, Grzeschik J, Krah S, Becker S, Könning D, Dickgießer S, Kolmar H: ***The shark strikes twice:*** **Hypervariable loop 2 of shark IgNAR antibody variable domains is capable of acting as an autonomous paratope.** *Mar Biotechnol (NY)* (2015), under revision

Zielonka S, Empting E, Kolmar H, Becker S, Hock B: **Method for generating bispecific shark variable antibody domains and use thereof.** *Patent application* (2014). Application no. EP 14003700.3

Zielonka S,[*] Empting M,[*] Grzeschik J, Könning D, Barelle CJ, Kolmar H: **Structural insights and biomedical potential of IgNAR scaffolds from sharks.** *mAbs* (2015) 7(1):15-25.

Zielonka S, Weber N, Becker S, Doerner A, Christmann A, Christmann C, Uth C, Fritz J, Schäfer E, Steinmann B, Empting M,Ockelmann P, Lierz M, Kolmar H: ***Shark attack*****: High affinity binding proteins derived from shark vNAR domains by stepwise in vitro affinity maturation.** *J Biotechnol* (2014) 191:236-245.

Doerner A,[*] Rhiel L,[*] Zielonka S,[*] Kolmar H: **Therapeutic antibody engineering by high efficiency cell screening.** *FEBS Lett* (2014) 588(2):278-287.

Further publications during PhD thesis:

Uth C, Zielonka S, Hörner S, Rasche N, Plog A, Orelma H, Avrutina O, Zhang K, Kolmar H: **A chemoenzymatic approach to protein immobilization onto crystalline cellulose nanoscaffolds.** *Angew Chem Int Ed Engl* (2014) 53(46):12618-12623.

[*] shared primary author

“I take the work seriously, just not myself in it.”

Henry Rollins

TABLE OF CONTENTS

1. ABSTRACT 1

1.1 Zusammenfassung 1

1.2 Abstract 2

2. INTRODUCTION 3

2.1 Antibodies as key players in the immune system 5

2.2 Structure and function of antibodies 5

2.3 Beneficial attributes and potential limitations of antibodies for biomedical applications 7

2.4 Bi-specific antibodies 8

2.5 The New Antigen Receptor (IgNAR) 9

2.5.1 Absence of a light chain partner 11

2.5.2 The variable domain of IgNAR – structural features 11

2.5.3 Diversification of the IgNAR repertoire 14

2.5.4 Selection of antigen-specific vNAR domains from shark immune repertoires 15

2.5.5 Therapeutic and diagnostic attributes of vNAR domains 16

2.5.6 Re-formatting of vNAR domains 18

2.6 Antibody engineering and selection 20

2.7 Target proteins 22

2.7.1 Epithelial cell adhesion molecule (EpCAM) 22

2.7.2 Receptor tyrosine kinase EphA2 22

2.7.3 Interleukin-8 (CXCL8) 23

2.7.4 Cluster of differentiation 3 (CD3) subunit CD3ε 23

2.7.5 Human Fcγ 24

2.8 Aim of the work 24

3. MATERIAL 26

3.1 Bacterial strains, yeast strains and cell lines 26

3.2 Plasmids 27

3.2.1 pCT 27

3.2.2 pMX-EETI 28
3.2.3 pMX-vNAR 29
3.2.4 pExpress-Fc 30
3.3 Enzymes, proteins, nucleic acids and protein ladder 31
3.4 Oligonucleotides 32
3.5 Chemicals 33
3.6 Solutions and buffers 36
3.7 Cell culture media 37
3.8 Laboratory materials and kits 38
3.9 Equipment 39
3.10 Animals 39

4. METHODS 40
4.1 Analysis of the natural repertoire of the bamboo shark 40
4.2 Yeast surface display library construction and establishment of randomized sublibraries for affinity maturation 41
4.2.1 CDR3-randomized library construction 41
4.2.2 Establishment of randomized sublibraries for affinity maturation 42
4.2.3 Construction of a HV2-diversified library for the isolation of bi-specific IgNAR V domains 42
4.3 Binding assays on the yeast surface and library screening for the isolation of target-specific vNAR molecules 43
4.3.1 Library analysis using yeast surface display 43
4.3.2 Single clone analysis using yeast surface display 43
4.3.3 Affinity titration on the yeast surface for the calculation of the equilibrium dissociation constant 43
4.3.4 Library screening using yeast surface display 44
4.4 Soluble expression of vNAR constructs in *E. coli* and protein purification 44
4.4.1 Cloning of vNAR fragments 44
4.4.2 Soluble expression and purification 45
4.5 Soluble expression of formatted vNAR-Fc fusion constructs in mammalian cells and protein purification 46
4.5.1 Formatting of vNAR fragments 46
4.5.2 Soluble expression of vNAR-Fc fusion proteins using the Expi293™ Expression System Kit 47
4.5.3 Soluble expression of vNAR-Fc fusion in HEK293 cells and purification 47

4.6 Characterization of vNAR proteins 48
4.6.1 Determination of binding kinetics using the Octet RED96 System 48
4.6.2 Thermal shift assays 49
4.6.3 Cellular immunofluorescence assays 49

5. RESULTS 50
5.1 Analysis of the natural vNAR repertoire of the bamboo shark (*C. plagiosum*) 50
5.2 Construction of a semi-synthetic type IV bamboo shark vNAR library for yeast surface display 51
5.3 Selection of antigen binding semi-synthetic vNAR domains 54
5.3.1 Screening against epithelial cell adhesion molecule (EpCAM) 54
5.3.2 Screening against receptor tyrosine kinase EphA2 56
5.3.3 Screening against human interleukin-8 58
5.4 Affinity maturation of target-enriched semi-synthetic vNAR domains 59
5.4.1 Sublibrary screening for the isolation of EpCAM binding vNARs with enhanced affinities 60
5.4.2 Sublibrary screening for the isolation of CXCL8 binding vNARs with enhanced affinities 61
5.5 Expression of selected vNAR domains 63
5.6 Formatting of α-EpCAM-vNAR 5005 as Fc-fusion 67
5.7 Characterization of selected vNAR domains 68
5.7.1 Affinity measurements using biolayer interferometry 68
5.7.2 Determination of thermal stability 70
5.7.3 Determination of selectivity with related and unrelated target proteins using yeast surface display 71
5.7.4 Tumor cell staining assays 73
5.8 Generation of a new paratope into the vNAR scaffold through randomization of HV2 74
5.8.1 Construction of a HV2 randomized library based on EpCAM-specific vNAR 5005 75
5.8.2 Selection of vNAR molecules targeting CD3ε 77
5.8.3 Selection of vNAR molecules targeting Fcγ 81

6. DISCUSSION 84
6.1 Library construction, selection of antigen-specific vNAR molecules and affinity maturation of enriched target binding variants 84
6.2 Generation of a new antigen-binding site into the vNAR scaffold through diversification of hypervariable loop 2 89

7. REFERENCES 92

8. APPENDIX 108

A Negative control of screening round 4 against receptor tyrosine kinase EphA2 108

B Multiple sequence alignment of vNAR molecules of different types from different species 108

C Affinities (K_D) determined by yeast surface display 109

D Abbreviations 111

E List of figures 113

F List of tables 114

G *Curriculum vitae* 115

9. AFFIRMATIONS 117

1. ABSTRACT

1.1 Zusammenfassung

Das Ziel der vorliegenden Arbeit war die Generierung von hoch-affinen Hai Antikörper vNAR Domänen, die verschiedene krankheitsrelevante Antigene adressieren. Hierfür wurde zunächst das natürliche IgNAR V Domänen Repertoire des Bambushais (*Chiloscyllium plagiosum*) charakterisiert und Kenntnisse hieraus dienten als Grundlage für die Etablierung einer Komplementaritäts-bestimmenden Region 3 (CDR3)-randomisierten Typ IV vNAR *Hefeoberflächendisplay* Bibliothek. Mittels Bibliothekendurchmusterung wurden verschiedene Antigen-bindende vNAR Domänen isoliert und hinsichtlich ihrer Affinität charakterisiert. Mit diesem Verfahren generierte Binder wiesen moderate Affinitäten, im dreistellig nanomolaren bis mikromolaren Bereich, für das jeweilige Antigen auf.

Für die Optimierung der Antigen-bindenden vNAR Domänen wurde eine neuartige Methode zur Affinitätsmaturierung etabliert, welche auf der Diversifikation der CDR1 der angereicherten Antigen-bindenden Varianten beruht. Durch die Konstruktion von Unterbibliotheken, in denen fünf Reste der CDR1-Schleife randomisiert wurden, sowie durch deren Durchmusterung mit signifikant reduzierten Antigenkonzentrationen wurden Affinitäts-verbesserte Klone identifiziert. Deren Affinitäten, bestimmt mittels *Hefedisplay*, waren im Vergleich zu den parentalen Molekülen, isoliert nach initialer Bibliothekendurchmusterung, erheblich optimiert.

Darüber hinaus wurden verschiedene vNAR Domänen als lösliche Proteine produziert und sorgfältiger hinsichtlich ihrer Affinität, mittels Biolayer Interferometrie, sowie in Bezug auf deren Stabilität, mittels *thermal shift assays*, charakterisiert. In diesem Zusammenhang ergab sich eine starke Korrelation zwischen den durch *Hefeoberflächendisplay* kalkulierten Affinitäten sowie den Affinitäten, bestimmt mit löslichen vNAR Domänen. Weiterhin wiesen alle produzierten vNAR Domänen eine hohe Thermostabilität auf. Ferner wurde eine EpCAM-bindende, affinitätsoptimierte IgNAR V Domäne als Fc-Fusion in Säugerzellen exprimiert. Durch die Charakterisierung dieser Variante mittels Biolayer Interferometrie konnte gezeigt werden, dass die Formatierung als Fc-Fusion zu einer moderaten, aber dennoch signifikant verbesserten Affinität um den Faktor von drei führt, vermutlich bewirkt durch Aviditätseffekte.

Des Weiteren wurde im Rahmen dieser Arbeit eine neue Antigenbindestelle in der variablen Domäne von IgNAR generiert. Dies wurde durch die Diversifizierung der hypervariablen Schleife 2 (HV2) der IgNAR V Domäne und Bibliothekendurchmusterung erreicht. Ein EpCAM-spezifischer vNAR wurde als Ausgangsmaterial genutzt und neun Aminosäurereste wurden in HV2 randomisiert. Zielmolekül-spezifische Klone, welche ein neues HV2-mediiertes Paratop beinhalteten, wurden gegen *cluster of differentiation 3* Untereinheit ε (CD3ε) sowie gegen den humanen Fcγ-Teil isoliert, wobei die hohe Affinität gegen EpCAM erhalten blieb. Folglich wurden somit bi-spezifische vNAR

Moleküle generiert. Im Wesentlichen konnte mittels *Hefeoberflächendisplay* gezeigt werden, dass ein neues Paratop in das vNAR Gerüst konstruiert werden konnte, welches unabhängig von der üblichen Antigenbindestelle, bestehend aus CDR3 und CDR1, fungiert.

1.2 Abstract

The aim of the work presented herein was the generation of high-affinity shark vNAR domains targeting different disease-related antigens. For this, the natural IgNAR V domain repertoire of the bamboo shark (*Chiloscyllium plagiosum*) was analyzed and in analogy to these findings, a semi-synthetic complementarity determining region 3 (CDR3)-randomized Type IV vNAR library was constructed for yeast surface display. By library screening against several disease-related antigens multiple different antigen binding vNAR domains were isolated and characterized in terms of affinity, revealing moderate affinities in the triple-digit nanomolar to micromolar range.

For optimization of antigen binding vNAR domains, a new methodology for affinity maturation was established that relies on the diversification of CDR1 of target-enriched binders. Sublibraries were constructed in which five residues of the CDR1 loop were randomized and affinity-enhanced vNAR domains were identified by sublibrary screening using significantly decreased target concentrations. Affinities determined using yeast surface display revealed substantially affinity-optimized clones compared to parental molecules, obtained from initial library sorting.

Additionally, several vNAR domains were produced as soluble proteins and characterized more meticulously in terms of affinity using bio-layer interferometry and in terms of stability *via* thermal shift assays. In this respect, affinities calculated by yeast surface display strongly correlated with affinities determined for soluble IgNAR V domains. Moreover, all produced vNAR variants exhibited high thermostability. Besides, an EpCAM-binding, affinity-matured vNAR domain was expressed as Fc-fusion protein in mammalian cells. Characterization of this formatted variant using biolayer interferometry resulted in a moderately, but significantly enhanced affinity by the factor of three, presumably through avidity-effects.

Furthermore, the generation of a new antigen-binding site into the IgNAR variable domain was also in the scope of this work. This was achieved through the diversification of hypervariable loop 2 (HV2) of the IgNAR V domain and library screening. An EpCAM-specific vNAR was used as starting material and nine residues in HV2 were randomized. Target-specific clones comprising a new HV2-mediated paratope were isolated against cluster of differentiation 3ε (CD3ε) and human Fcγ while retaining high affinity for EpCAM, resulting in bi-specific vNAR molecules. Essentially, it was demonstrated that a new paratope can be engineered into the vNAR scaffold that acts independently from the original antigen-binding site, composed of CDR3 and CDR1, as verified by yeast surface display.

2. INTRODUCTION

Vertebrates, invertebrates as well as all multicellular organisms are constantly exposed to microbial pathogens and potentially life-threatening infections. However, those hosts are usually capable of defending themselves against pathogenic substances by their ability to recognize infectious microbes and *via* inducing an adequate defense response.[1,2] The evolutionary ancient form of defense mechanisms shared by most - if not all - multicellular organisms is referred to as innate immunity. In addition, jawed and jawless vertebrates possess highly specialized and antigen-specific defense mechanisms in order to overcome exposure to severe pathogens. This kind of defense is called adaptive immunity. The innate immune systems as well as adaptive immunity are as extensively described by Medzhitov and co-workers [3,4]

Innate immunity can further be subdivided into constitutive and inducible defense mechanisms.[1] Constitutive defense mechanisms e.g. barrier functions, antimicrobial peptides and enzymes are present at sites of continuous interactions with microbes and their defense responses are strictly directed to target potential pathogens but not host cells. Inducible defense requires recognition of the invader and ultimately results in triggering cascades activating inflammatory responses and killing of the potential pathogen.[5,6] Key elements for the recognition of invading organisms are so called pathogen recognition receptors (PRRs). These receptors can either be expressed on the cell surface, in intracellular compartments or can be secreted.[7] PRRs recognize structural elements only produced by microbial pathogens, referred to as pathogen-associated molecular patterns (PAMPs). PAMPs represent conserved molecular patterns, often shared by large groups of microorganisms.[1,8] Hallmarks of PRRs are that those molecules are usually fixed in the genome, i.e. no rearrangement events occur to produce a very diverse set of receptors. Furthermore, there is no clonal distribution, hence, all cells of a particular class usually express identical PRRs. Additionally, PRRs are present before the exposure with the invader. As a consequence, effector functions are immediately activated upon induction of the receptor.[5]

Central elements of innate immune mechanisms are present in invertebrates and also in vertebrates. However, it is supposed that the variability and diversity is much higher among invertebrates, which in contrast to vertebrates are unable to respond in an antigen-specific manner.[6] Essentially, according to Flajnik and Du Pasquier, the invertebrate innate immune system varies greatly in the level of complexity and in its forms of defense among different organisms and invertebrates immunity evolves rapidly.[9] Within the animal kingdom several hosts have evolved their own special mechanisms of defense and it is not always that easy to clearly discriminate between innate and adaptive immunity.[9-11]

One striking example is the hypervariable gene that encodes for Dscam molecules (Down syndrome cell adhesion molecule) in insects, which is involved in the innate immune system.[12] A diverse set of

more than 1.9 x 10^4 different Dscam molecules with a molecular weight of approximately 210 kDa can be generated by alternative splicing. Dscams function as PRRs and it was shown that a knock-down of Dscam expression impairs bacterial uptake by insects.[10-12]

Another peculiarity is found in snails which express a very diverse family of fibrinogen-related proteins (FREPs).[13] FREPs consist of one or two *N*-terminal Ig superfamily domains followed by a *C*-terminal fibrinogen domain. FREP gene expression is up-regulated following infection and FREP proteins are able to bind a wide range of pathogens displaying carbohydrates on their surfaces. FREPs mediate several different immune effector functions e.g. agglutination, phagocytosis and the release of toxic oxygen radicals. [11] FREP genes diversify somatically by gene conversion and/or somatic hypermutation as well as alternative splicing.[9] Interestingly, as described by Flajnik und Du Pasquier, somatic mutations are associated with adaptive immunity which is exclusively ascribed to vertebrates.[9]

Hallmarks of the adaptive immune response are antigen-specific responses and memory. Key elements of the jawed vertebrate's adaptive immune system are B-cell receptors (BCRs), T-cell receptors (TCRs) and MHC molecules.[14] However, in contrast to this, the jawless vertebrate's (lampreys and hagfish) acquired immunity relies on a unique class of antigen receptors referred to as variable lymphocyte receptors (VLRs).[15,16] VLRs are structurally unrelated to BCRs and TCRs and are composed of highly diverse leucine-rich repeats (LRRs) sandwiched between amino- and carboxy-terminal constant LRRs. There are three different types of VLRs, called VLRA, VLRB and VLRC. VLRA and VLRC are expressed on T-cell like VLRA+ and VLRC+ cells. VLRB+ cells expressing VLRB are B-cell like and differentiate into VLRB secreting plasma cells after antigen stimulation.[14] The theoretical diversity of the VLRA and VLRB repertoire is estimated to vary between 10^{14} and 10^{17} different receptors, respectively, thus it is comparable to the potential diversity of TCR and BCR repertoires in mammals.[17] The jawless adaptive immune system and structural as well as functional features of VLRs are elegantly described elsewhere.[14,17]

In addition to the above described peculiarities, the jawed vertebrate's immune system also comprises several peculiar features e.g. heavy-chain only antibodies or antibodies with ultra-long CDR3-regions, some of which will be discussed in this work.[18,19]

Typically antibody molecules are composed of heavy chains and light chains. However, camelids and the cartilaginous fish possess natural antibodies composed only of heavy chains (HCAbs).[20,21] The antigen-binding sites i.e. paratopes of those exceptional molecules are formed by only one single domain, referred to as VHH and vNAR, respectively. Due to an increased frequency of polar and charged amino acids at the solvent-exposed regions corresponding to the hydrophobic VH-VL interface of conventional antibodies, vNAR and VHH domains are highly soluble.[22] Intriguingly, HCAbs naturally complement the conventional repertoire of the aforementioned species. Whereas classical antibodies usually have planar or concave antigen-binding sites, vNAR- and VHH-domains

possess a wide variety of different loop structures. This leads to a drastically expanded repertoire of available paratopes capable of accessing and binding to more cryptic epitopes and catalytic clefts of enzymes that are intractable to classical antibodies.[23-25] Due to this paramount therapeutic and diagnostic attribute of VHHs and vNARs, there is a strong effort within the scientific community to utilize these domains for medical and biotechnological applications. While camelid VHH domains have proven to be successful in early phase clinical trials,[26] the engineering of vNAR domains - which are the main focus of this work - for biomedical applications is at an earlier stage.

2.1 Antibodies as key players in the immune system

Antibodies are part of the adaptive immune system and comprise the humoral immune response. At first, antibody molecules are expressed as receptors on the surface of B-cells. It is supposed that the immune system gives rise to theoretically more than 10^{14} different B-cells, each producing a unique surface bound antibody (B-cell receptor) with a distinct specificity.[27] The process for the primary diversification of the antibody repertoire will be shortly discussed in section 2.5.3. Initial interaction of a B-cell receptor that specifically binds to the particular antigen provokes a primary response. As a consequence and also owing to complex immunological processes that involve parts of the innate and of the adaptive immune system, the B-cell starts to secrete the antibody molecule.[2-4] The high specificity for the individual antigen combined with Fc-mediated immune effector functions of the antibody ultimately causes the elimination of the antigen. After further stimulation with the antigen the secondary antibody response is characterized by high-affinity antibodies of other classes, mediated by processes referred to as somatic hypermutation and isotype switching, which will not be discussed within the scope of this work.[28-30] In humans there are five isotypes of antibodies categorized based on their heavy chain constant domains, IgM, IgG, IgA, IgD and IgE. In addition IgG can be divided into four subclasses, namely IgG1, IgG2, IgG3 and IgG4 and akin to this IgA can be subdivided into IgA1 and IgA2.[31] Essentially, the isotype switch from IgM to another isotype alters effector functions,[31,32] which will be discussed in the next chapter. Furthermore the isotype determines the pharmacokinetic profile of the antibody as well as the localization of the antibody within the body.[32] Because most of the therapeutic antibodies in the clinic are of the IgG isotype,[33,34] this class will be described more in detail.

2.2 Structure and function of antibodies

IgG antibodies are structurally complex, large hetero-tetrameric proteins (approximately 150 kDa) comprising two identical heavy chains and two identical light chains (**Fig. 1**). Each heavy chain consists of three constant domains, followed by the *N*-terminal variable domain. The light chain is composed of one constant domain and one *N*-terminal variable domain. The two identical antigen-

binding sites, i.e. paratopes, are composed of one variable domain of the heavy chain and one variable domain of the light chain, respectively.[34] Within each variable domain, three hypervariable loops, called complementarity determining regions (CDRs), mediate antigen binding.[31] Hence, a total number of six CDRs (three of the heavy chain and of the light chain, respectively) may contribute to antigen binding. Antibodies bind antigens *via* non-covalent interactions. Antibody molecules can act directly on targeted antigens, e.g. the antigen-antibody interaction can neutralize the function of the antigen or block downstream-signaling of a cell-surface-receptor on a targeted cell.[35-37]

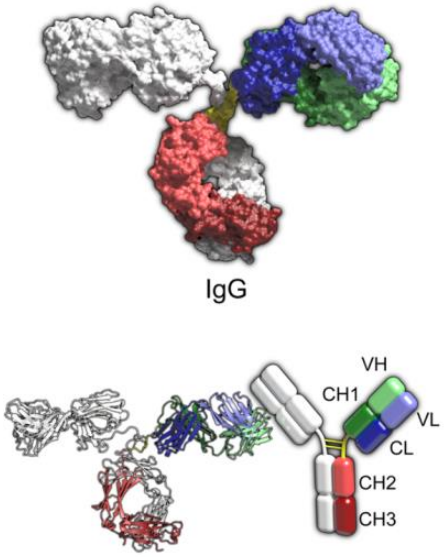


Figure 1. Structural features of IgG antibody shown as surface representation (top) as well as ribbon and schematic representation (bottom). Individual domains are colored as indicated in the schematic representation; hinge region is colored yellow; glycans not shown. IgG model is based on pdb entry 1IGT.[38] Picture rendered with POV-Ray (www.povray.org/). Figures 1 – 6 were constructed in close cooperation with Martin Empting and have recently been published.[39]

In addition, antibodies are able to act indirectly *via* effector functions. Such functions are mediated by the Fc-part of the molecule (i.e. CH2 and CH3 of each chain of IgG). Binding to Fcγ receptors, for instance, that are expressed on almost all cells of the immune system, can result in activation of cells and effector mechanisms such as antibody-dependent cellular cytotoxicity (ADCC) mediated by natural killer cells, phagocytosis of immune complexes or the release of inflammatory mediators.[32,40] Essentially, Fc receptors link the humoral components of the adaptive immunity with cellular components of immunity. On the same cells, there are activating and inhibitory receptors co-expressed. Balanced signaling through these keeps the immune response in normal limits.[41] The main interaction site of IgG with Fcγ receptors is located in the lower hinge-region (**Fig. 1**, yellow) and in the adjacent sites of CH2 (**Fig. 1**, light red).[32,42] Moreover, a glycosylated Asn-297 in CH2

of IgG is critical for Fcγ receptor binding. Finally, also residues of CH3 are important for this interaction.[32]

Another receptor, referred to as neonatal Fc receptor (FcRn) is responsible for the delivery of IgG across the maternofetal barrier and also for controlling the catabolism of this isotype. The mechanism that ensures the integrity of both processes is the ability of IgG to interact with FcRn in a pH-dependent manner. IgG is bound by FcRn after internalization in the endosomal compartment at mild acidic conditions (pH 6.0 – 6.5) and released at pH 7.0 – 7.4 on the plasma membrane and thus recycled as well as transported within and across cells. [32,43] Essential for FcRn interaction, hence, for transcytosis as well as IgG-recycling are residues located at the interface of domains CH2 and CH3 of the Fc-part. It needs to be mentioned that there are also proteins and receptors that interact with Fc-parts of other antibody isotypes. Such interactions can also mediate effector functions or mediate the transport of IgA and IgM.[32]

In addition to FcR-binding, IgG molecules (as well as other isotypes) interact with several other proteins e.g. components of the complement systems (C1q, C3b, C4b) and complement-related anaphylotoxins. Herein only the interaction with C1q is described. A comprehensive review about Fc-mediated interactions and effector fuctions can be found elsewhere.[32] IgG as well as IgM molecules are able to induce the classical pathway of the complement system by their ability to bind to C1q. As monomers IgG molecules display very weak affinity to C1q and are consequently unable to activate the complement cascade, which ultimately leads to the formation of the membrane attack complex and the elimination of the targeted cell. This process is named complement-dependent cytotoxicity (CDC). However, owing to polyvalent interactions, IgG aggregation i.e. multivalent interactions of several IgG molecules increases apparent affinity significantly, resulting in the activation of the complement cascade. C1q binding to IgG is primarily due to structural features of CH2 of IgG and multiple residues are involved in the interaction.[32,44]

2.3 Beneficial attributes and potential limitations of antibodies for biomedical applications

Parts of this section were recently published.[39] Nowadays, biological entities are one of the main drivers of sales within the pharmaceutical market, with growth rates that substantially succeed those of the overall pharmaceutical sector [45]. Within this group of new biological entities, monoclonal antibodies (mAbs) are among the most successful and frequently used biotherapeutics. mAbs are the highest selling class of biologics, followed by hormones, growth factors and fusion proteins, respectively.[45] Annual growth rates of mAbs are constantly increasing within the last years,[45-47] exceeding 10 % for the first time in 2011 [46] and forecasts predict an ongoing trend.[48] The total number of 40 US Food and Drug Administration (FDA) - approved antibodies with several

blockbuster agents amongst them [48] and a continuously growing number of more than several hundred of mAbs in clinical development is clearly emphasizing their paramount therapeutic value.[45,49,50]

However, under certain circumstances the therapeutic and diagnostic efficacy of monoclonal antibodies might be limited due to inherent attributes, e.g. structural complexity and large size. The paratope of conventional antibodies can be restricted in its ability to access certain epitopes, e.g. recessed cryptic epitopes or active sites of enzymes.[18,51,52] Furthermore, the mobility, i.e. tissue penetration, of classical antibody molecules is constrained by their large size.[53] For *in vivo* tumor imaging purposes, the slow blood clearance of conventional antibodies poses a problem due to their extended plasma half-life.[54] Slow tumor penetration as well as nonspecific uptake by healthy tissues may represent further drawbacks of conventional antibodies in molecular imaging.[55-58] Moreover, most complex diseases, e.g. cancer, HIV and inflammatory diseases are usually multifactorial in nature. Thus, tumor cells often up-regulate different growth-promoting receptors that can act either independently or crosstalk intra-cellularly.[59] Targeting of one receptor by a mono-specific antibody may result in resistance which for instance may be associated with the up-regulation of alternative receptors as well as pathway switching.[60]

To address these issues and to increase the overall therapeutic and diagnostic efficacy, next-generation-antibodies (e.g. bi-specific antibodies, antibody-drug-conjugates, antibody-cytokine-conjugates), antibody fragments and non-immunoglobulin based protein scaffolds have been engineered and developed, as extensively described elsewhere.[61-70] Because bi-specific antibodies are of particular relevance for this work, the rationale for their design as well as different formats will be discussed in the following section.

2.4 Bi-specific antibodies

There is a plethora of rationale for the generation of bi-specific antibody molecules (bsAbs) simultaneously binding two different antigens with their two antigen-binding sites. As aforementioned, many diseases are multifaceted in their behavior.[59,71] Crosstalk of receptor tyrosine kinases, for instance, can induce tumor resistance to (mono-) therapy. For example, there is clear evidence that tumor resistance to an inhibitor, i.e. Cetuximab of the Epidermal Growth Factor Receptor (EGFR) is caused by the up-regulation of other human epidermal growth factor receptor (HER) members.[72,73] Furthermore crosstalk between insulin-like growth factor receptor 1 (IGF-1R) and EGFR may mediate resistance to one another.[72,74,75] Besides, also overexpression as well as activation of the hepatocyte growth factor receptor (cMET) is known to induce tumor resistance to HER member blockage.[73] bsAbs can overcome the limitation of mono-specific antibody based therapy and find application in multifactorial diseases because those molecules can be engineered to address multiple

epitopes of the same target or sites on different targets within one single entity.[71,76,77] In addition, selective targeting of a defined cell population characterized by the overexpression of two antigens can improve safety (especially for bi-specific antibody-drug conjugates) and efficacy. Other approaches for the design of bi-specific antibodies comprise the transport of the drug to a specific tissue e.g. delivery across the blood-brain barrier as well as the recruitment of effector cells.[78] In the latter respect, targeting a tumor-associated antigen as well as CD3 of cytotoxic T-lymphocytes can elicit effective tumor killing.[79] Besides, targeting specific Fcγ receptors can engage natural killer cells and myeloid effector cells to eliminate tumor cells.[80] Another rationale for bsAbs is the utilization of such molecules for the delivery of payloads such as imaging agents for diagnostic applications. In this pre-targeting application, one antigen-binding site of the bi-specific molecule interacts with the clinical target, the other arm is able to bind to a detection moiety, that is injected once the unbound bsAb has cleared from the blood-stream.[71,80]

The first bsAb that has been approved for therapy in the European Union in April 2009 was Cetumaxomab for the treatment of patients with malignant ascites.[81] Cetumaxomab is a bsAb, simultaneously targeting the epithelial cell adhesion molecule (EpCAM) which is expressed on a multitude of tumors and CD3 on T-cells for the redirection and activation of T-cells at the tumor site resulting in significant tumor lysis. This therapeutic entity is a combination of half of an EpCAM-specific mouse IgG2a and a half of a CD3-specific rat IgG2b.[82]

To this date, there is a multitude of different bsAb formats targeting multiple different antigens in clinical development.[79,80] bsAb formats include but are not limited to trifunctional hybrid antibodies (e.g. Cetumaxomab), tandem scFvs such as bi-specific T-cell engagers (BiTEs), bispecific diabodies as well as the Dock and Lock (DNL) platform.[79,82] Moreover, different methodologies have been established to design human full-length IgG-like antibodies such as knob-into-holes[83], SEED[84] or Luz-Y[85] for the hetero-dimerization of the heavy chain as well as CrossMab technology for correct association of a respective light chain to its corresponding heavy chain.[86] Essentially, bsAbs represent a whole new repertoire for therapeutic applications and in contrast to antibody mixtures there is no need for time-consuming and costly separate single agents trials followed by combination trials.[78,82]

2.5 The New Antigen Receptor (IgNAR)

A modified version of this chapter was recently published in a review on shark antibodies.[39] The cartilaginous fish (sharks, rays, skates and chimaeras) express three different isotypes of antibodies, IgM, IgNAR and IgW.[87,88] IgNAR was first identified in the serum of the nurse shark (*Ginglymostoma cirratum*) in 1995 by Flajnik and co-workers.[21] It is a homodimer of heavy chains devoid of light chains. Each chain of the secretory form consists of one variable domain followed by

five constant domains, the last four being homologous to IgW constant domains.[89] Serum IgNAR levels range from ~ 0.1 mg/ml to 1 mg/ml.[90]

Based on atomic resolution structural data as well as small-angle X-ray scattering, Buchner and co-workers were able to develop a structural model of the complete IgNAR molecule (**Fig. 2**).[91] IgNAR is a heavy-chain only homodimer. Each heavy chain consists of five constant domains, followed by the *N*-terminal variable domain. Within the molecule, domains C1 and C3 of each chain cause dimerization of IgNAR. Despite the lack of a canonical hinge region, the variable domains are spaced sufficiently wide for binding multiple epitopes, facilitated by the wide angle of the C1 dimerization interface. A small angle between both C3 domains induces the formation of a narrow stalk for the IgNAR molecule. However, flexibility of the stalk is induced by a disulfide-bridged linker that connects domains C3 and C4. The heavy-chain only molecule is kinked approximately in the middle of the molecule, at the location of the flexible linker, causing its characteristic shape. Whether any effector functions are mediated by the constant region of IgNAR is currently unresolved.[51] Finally, it is important to note that the structure of C5 shown in **Figure 2** is completely hypothetical (**Fig. 2**). This is due to a lack of structural data on this domain, which does not exhibit a fold, neither as isolated recombinant protein nor within a C4-C5 construct.[91]

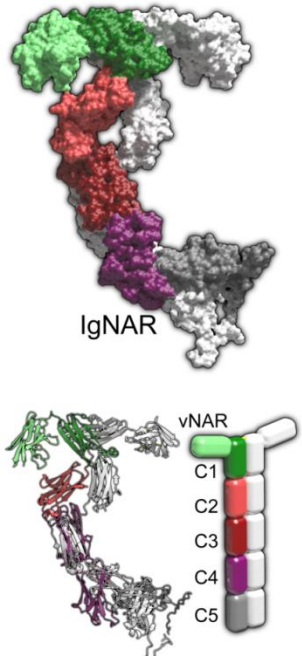


Figure 2. Structural features of shark IgNAR antibody formats shown as surface representation (Top) as well as ribbon and schematic representations (Bottom). Individual domains are colored as indicated in the schematic representation. Coordinates of intact IgNAR including the hypothetical structure of IgNAR C5 domain were generously provided by Prof. Dr. Michael Sattler and Dr. Janosch Hennig (see Feige *et al.*).[91] Picture rendered with POV-Ray (www.povray.org/).

2.5.1 Absence of a light chain partner

The homodimer IgNAR displays several unique features that are responsible for the inhibition of a potential light chain pairing. At the typical VH-VL interaction site, there is poor conservation of residues that mediate this association in mammals.[92] Instead these typically hydrophobic amino acids are frequently replaced by polar or charged residues.[22] For classical antibodies, a special mechanism ensures the formation of heavy- and light-chain pairing. In the endoplasmic reticulum, the heavy chain is trapped by an Ig-binding protein (BiP) *via* interaction with the CH1 domain. For the release, a light chain must displace BiP, and, consequently, only heavy- and light-chain paired antibodies are secreted.[22,93] Flajnik and co-workers hypothesized that during evolution, a vNAR-D-J cluster recombined with an IgW cluster in a way that the IgW cluster lost its V-D-J segments and the first C exon.[94] Indeed, the C1 domain of IgNAR is somewhat similar to the CH2 domain of IgW and may be derived from this domain.[95] BiP- and L-chain-interaction sites are consistently missing in the C1 domain of IgNAR, as elegantly reviewed by Flajnik and colleagues.[22]

2.5.2 The variable domain of IgNAR – structural features

The variable domain of the New Antigen Receptor shows homology to the T-cell receptor (TCR) V_α and also is found as a variable domain in the NAR-TCRδ.[96] It also displays sequence homology to immunoglobulin V_κ domains, whereas structurally it is related to V_α, V_λ, and VH domains.[24] Moreover, since vNAR domains share structural features of cell adhesion molecules, it was suggested that IgNAR evolved from a cell-surface receptor, clearly distinguishing it from VHH, which evidently arose from an IgG lineage.[22,97] vNAR belongs to the Ig superfamily, and accordingly it has a β-sandwich fold. However, compared to mammalian V domains, this fold only consists of 8 instead of 10 β-strands due to the deletion in the framework2-CDR2-region (**Fig. 3**).

With a molecular mass of ~ 12 kDa, the vNAR domain is the smallest antibody-like antigen binding domain in the animal kingdom known to date.[24,51] As a consequence, contrary to mammalian variable domains, vNAR domains have only two complementarity determining regions CDR1 and CDR3 (**Fig. 3, Fig. 4**). The diversity of the primary vNAR repertoire is predominantly found in CDR3. High rates of somatic mutation after antigen contact are observed in CDR1, at the CDR2 truncation site, where the remaining loop forms a belt-like structure at the bottom of the molecule and in a loop that corresponds to HV4 in TCRs. Accordingly, these mutation-prone regions have been named HV2 and HV4, respectively (**Fig. 4**).[98] Indeed, it was shown that somatic mutations within HV4 can contribute to antigen binding.[99]

Despite having a reduced number of possible antigen binding loops (four across a single chain) compared to conventional antibodies (six loops across two chains), vNAR domains bind antigens with surprisingly high affinities.[100,101] Even from primary repertoires, where antigen binding is solely

mediated by CDR3, vNAR molecules can be obtained against a given antigen with affinities in the low nanomolar range.[99,100] The highest recorded affinities for vNAR domains, however, have been observed after immunization with an anti-albumin binding domain, achieving picomolar levels of affinity.[101]

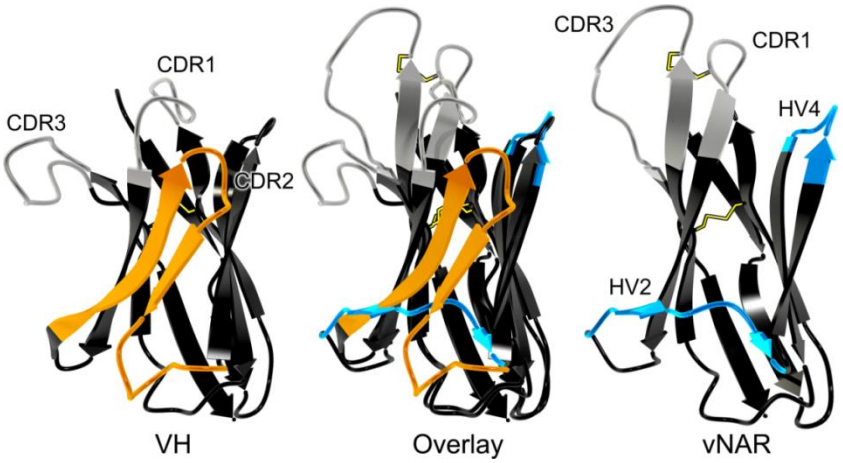


Figure 3. Comparison of VH (left; from pdb entry 1IGT) and vNAR (right, from pdb entry 2COQ) binding domains depicted as ribbon representation as well as an overlay of both structures (middle).[25,38] CDR1 and CDR3 are shown in grey. Two beta strands and CDR2 of the VH domain are highlighted in orange. These structural elements are absent in the vNAR domain which possesses HV2 and HV4 (both highlighted in blue), instead. Disulfide bonds are shown as yellow sticks. Picture rendered with POV-Ray (www.povray.org/).

Based on the number of non-canonical cysteine residues, which are not found in classical variable domains, vNAR molecules have been categorized into four types (**Fig. 4**).[24,25,99,102,103] The classical Ig canonical cysteines, which stabilize the immunoglobulin fold *via* a disulfide bond, are common to all types. Type I variable domains carry extra cysteines in framework regions 2 and 4, and consequently, an even number of partner cysteine residues in CDR3. The determination of the crystal structure of a type I vNAR in complex with lysozyme revealed that both non-canonical framework cysteines each form disulfide bonds with those of CDR3, causing this loop to be held tightly into the direction of HV2.[24] Thus far, type I variable domains of IgNAR have only been identified in the nurse shark, *Ginglymostoma cirratum*.[51]

Type II domains differ from type I by means of an additional cysteine in CDR1 and in CDR3, respectively, resulting in an intra-molecular disulfide bond that brings both loops in close vicinity. However, it lacks both cysteine motifs that anchor CDR3 to the framework in type I vNAR. As a consequence, the CDR3 region forms a protrusive ‘finger-like’ structure that is predisposed to binding into pockets or grooves, e.g. the active site of enzymes.[25] According to this, it has been shown for both types that active site clefts can be penetrated by the antigen binding loops.[24,99]

Another type, termed type III, is expressed in neonates at high frequencies.[25,102] Akin to type II domains, this isotype is characterized by an additional non-canonical cysteine in CDR1 and CDR3, respectively. However, in contrast to type II, type III domains comprise a restricted CDR3 diversity

that is highly similar in amino acid composition and length as well as a conserved tryptophan residue in CDR1 positioned adjacent to the disulfide bridge between both loops. Based on the limited CDR3 diversity it is tempting to speculate that type III vNARs evolved as a consequence of exposure to a common pathogen in early development of sharks or that it may play a role in regulatory processes during the development of the shark's immune system.[25,51,102]

Type IV domains differ from all described vNAR types in that they lack non-canonical disulfide bonds as found in all other vNAR types.[101,103] Therefore, the topology of the paratope of type IV variable domains is more flexible and not physically constrained. Type IV domains are also referred to as type IIb, according to Streltsov *et al*. and Liu *et al*.[97,104] In addition, type IV domains with an invariant tryptophan residue in CDR1, similar to type III, have been identified.[105] Besides type III, all types of the vNAR domain give rise to high-affinity binders.[24,99,101,106]

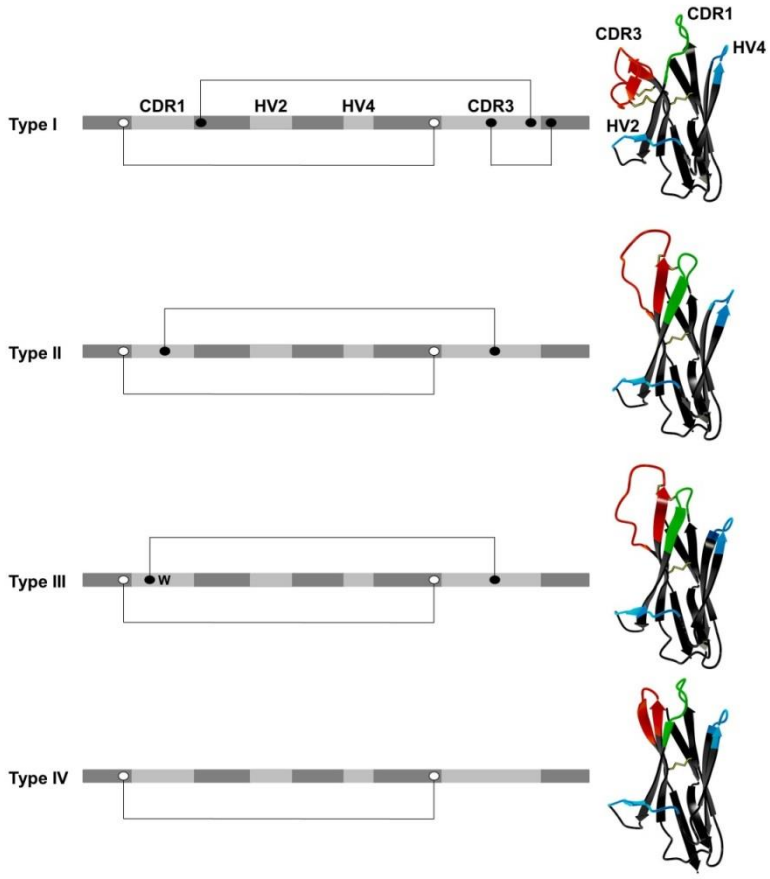


Figure 4. Different types of IgNAR V domains. Variable domains are categorized based on the presence or the absence of non-canonical cysteine residues (black dots). Canonical cysteine residues (white dots) and disulfide bonds (connecting lines), conserved tryptophan (W) as well as complementary determining regions (CDR) and hypervariable loops (HV) are shown in their relative positions. Ribbon presentations of vNAR domains are depictions of pdb entries 1SQ22 (type I),[24] 2COQ (type II),[25] and 4HGK (type IV)[103] as well as a modelled type III structure based on 2COQ. The latter was generated *via* homology modeling using YASARA structure.[107] First, vNAR residues of 2COQ were changed to match a reported type III sequence (AAM76948 from Streltsov *et al.*).[25] Then, side chain geometries were optimized followed by a two-step energy minimization using the YASARA2 force field.[107]

2.5.3 Diversification of the IgNAR repertoire

In mammals, antibody diversity is generated in a process referred to as V(D)J-recombination. During B-cell development, one variable (V) segment, one diversity (D) segment and one joining (J) segment are randomly rearranged from a multitude of gene segments of the immunoglobulin heavy chain gene cluster to encode the VH domain, which is fused to a gene segment encoding a constant domain.[22] Similarly, for light chain generation, one V segment and one J segment are selected by chance from a pool of gene segments to produce the variable domain of the light chain that is fused to a CL gene. Diversity is further expanded in a process called junctional diversification through imprecise segment joining.[108] An additional layer of diversity is introduced by the random arrangement of the heavy chain and light chain to complete the expression of the antibody molecule.

As IgNARs are devoid of light chains, they lack H-L combinatorial diversification. Correspondingly, one would expect a dramatically restricted primary repertoire. However, this lack of diversification process is at least partially compensated through the diversity achieved within the CDR3 region. Whereas mammalian antibody genes are organized in the translocon-format, shark antibody genes are exclusively arranged in the cluster-organization (**Fig. 5**).[24,51,94,99] Nurse sharks comprise four IgNAR gene clusters, though only two are expressed in adult life, one encodes type I and one encodes type II IgNARs.[24,99] Each IgNAR cluster comprises one V segment, three D segments and one J segment and a single set of C segments. Rearrangement occurs solely within this cluster resulting in a $VD_1D_2D_3J$ assembly. Hence, four rearrangement processes generate the complete vNAR domain. However, the order of rearrangements remains to be determined.[94] The interfaces between the V segment, the three D segments and the J segment encode for CDR3. Consequently, diversity in both sequence and length of the primary repertoire is nearly entirely found in CDR3.[102,109,110] Extensive junctional diversification through N-region addition, P-nucleotide addition, trimming and D-region rearrangement further expands the heterogeneity of the primary repertoire of IgNAR.[24,51,94,111] The type III gene cluster represents an exception within the recombination process. As a result of the fusion of the D_1 segment and the D_2 segment in the nurse shark, only three rearrangement events occur, explaining the restricted diversity of this type. In contrast to this, in the spiny dogfish the type III IgNAR cluster is not partially germline-joined, indicating that germline-joining of Ig clusters might be a species-specific event.[112]

Sharks do not possess conventional germinal centers. Nevertheless, the initial combinatorial diversity which is mainly restricted to CDR3, is further expanded by extensive somatic hypermutation in an antigen-driven manner, with mutations clustering to the CDRs.[109] The mutational pattern and frequencies of this process are similar to that of mammalian immunoglobulins with a bias for transitions over transversions. The mechanism favors the serine codon AGC/T as a hotspot for mutations and most of the changes are base substitutions. Surprisingly, base changes often occur in tandem, particularly in mutational hotspots and palindromic repeats.[110] It was first shown by Flajnik

and co-workers, that after immunization, somatic mutations promoted an incremental increase in affinity, giving clear evidence for *in vivo* affinity maturation in sharks.[24,90,98-100] Furthermore, they demonstrated that HV4 is prone to somatic mutations and, even more interesting, these mutations can be involved in antigen binding.[99]

Consistent with the structure of type I and type II IgNARs (**Fig. 5**), mutations are favored in CDR1 for type II vNARs and in HV2 for type I vNARs.[22,24] It is hypothesized that those mutations could either directly contribute to antigen binding or they could indirectly have an effect on the paratope such as that they stabilize and influence the conformation of the antigen binding CDR3.

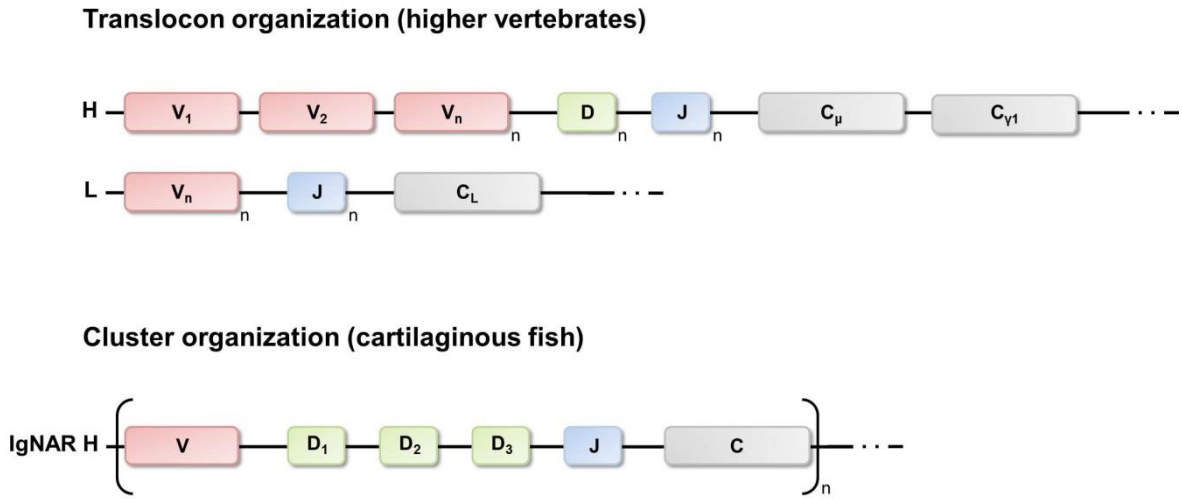


Figure 5. Translocon arrangements of immunoglobulin genes in higher vertebrates and cluster configuration of IgNAR genes of cartilaginous fish. In the translocon organization there are many variable (V) segments upstream of many diversity (D, only for heavy chains) and joining (J) segments that recombine randomly to encode the variable domain of the heavy chain or the light chain. IgNAR genes (like all Ig genes of the cartilaginous fish) are organized in the cluster configuration. Each cluster contains for the variable domain one V segment, three D segments and one J segment. Recombination occurs exclusively within one cluster. H, heavy chain loci; L, light chain loci, C, constant region.

2.5.4 Selection of antigen-specific vNAR domains from shark immune repertoires

Antigen-specific vNAR domains have been generated from the immune repertoire of a number of different shark species, including the nurse shark,[100] the wobbegong shark,[106,113] the spiny dogfish,[101,104,114] and the banded houndshark.[115,116] Target-specific clones are generally isolated using different display technologies, such as phage display[100,106] or ribosome display.[117] There are several distinct strategies for library establishment. Binders can be selected from immunized sharks,[100,101] from the naïve shark repertoire,[104] or from a synthetic vNAR library,[106,118] where the vNAR molecule serves as a scaffold with randomized loops and from semi-synthetic repertoires. Here, additional diversity is included through the randomization of one or more antigen binding loops.[114,119] For the most part, immunization is the preferred route to obtain high-affinity binders. An additional advantage of immunization is that sharks are evolutionary very distant to humans. This greatly reduces the likelihood of immune tolerance that would reduce the

induction of target-specific responses when antigens are well conserved across species. Consequently, antigen-specific vNAR molecules have been generated with impressive affinities against a multitude of different targets *via* immunization.[100,101,120,121] However, the process of immunization of sharks is protracted compared to standard mammalian protocols, and not every species tested has proven successful.[101,122] For instance Dooley *et al*. and also our group were unable to detect an antigen-specific IgNAR response after the immunization of the small spotted catshark, *Scyliorhinus canicula.*[123]

Antigen-specific vNAR fragments have also been isolated from non-immunized libraries against a plethora of different targets, including viral targets and toxins.[104,114,116,119,124,125] **Table 1** illustrates vNARs selected against therapeutically relevant targets from immunized and non-immunized origins (**Tab. 1**). Surprisingly, binders selected from such libraries often show good affinities to their target.[113,114,124,125] Nonetheless, when higher affinities are required, vNARs can be optimized using *in vitro* affinity maturation. To this end, two different methodologies have been established. Nuttall and co-workers were able to improve the affinity for an AMA1-specific IgNAR V domain using error-prone PCR resulting in a ~ 10-fold enhanced affinity.[125] In a more recent approach, the same group employed a mutagenesis system dependent on low fidelity RNA polymerase from Qβ bacteriophage to introduce diversity into IgNAR antibody libraries for affinity maturation. With this novel strategy they were able to select mutated vNAR molecules with a more than 20-fold enhanced affinity compared to the wild type clone.[117]

2.5.5 Therapeutic and diagnostic attributes of vNAR domains

The tremendous diversity found at the sequence-level of the CDR3-loop of IgNAR, as well as the multiplicity of the structural topologies formed by the antigen-binding site of the vNAR domain (**Fig. 6**) render IgNARs promising alternatives to conventional antibodies.[24,25,99] As described above, the different types of vNAR domains form, if any, a very diverse set of disulfide bridges. Consequently, antigen-specific clones can be selected from a very large, unprecedented repertoire of different loop structures.[126] Moreover, this unique paratope-architecture of shark domains seems to be predisposed to target clefts of the antigen, whereas recessed epitopes are usually not antigenic for conventional antibodies.[22,25,97,127] Indeed, it has been shown that the active site of enzymes and clefts can be targeted by vNAR domains.[24,91,99]

Above all, vNARs exhibit many additional properties that render them interesting for diagnostic and therapeutic applications. It has been demonstrated that vNARs are extraordinarily stable proteins,[100,104,114,118,121] which is probably a consequence of the harsh physiological environment - the blood of sharks contains 350 mM urea - those molecules are exposed to.[105] The superior thermal stability and tolerance to irreversible thermal denaturation compared to scFv- and

mAb-formats was elegantly demonstrated by Lonsdale and colleagues as well as by Goldman and co-workers.[104,128] Concordantly, it has been shown that the C2 and C4 domains of IgNAR are very stable. Buchner and co-workers were able to identify structural elements that contribute to their high stability. Compared to mammalian constant domains, C2 and C4 domains of IgNAR contain an additional salt bridge and an extended hydrophobic core. The transfer of these key elements of enhanced stability to a human antibody domain improved its stability significantly.[91]

The inherent small size of the IgNAR V domain is an additional therapeutic and diagnostic attribute. It can be hypothesized that this property leads to a greater mobility with reference to tissue penetration. Especially for *in vivo* imaging, where a high contrast to background ratio is crucial, this feature is beneficial, due to an advantageous pharmacokinetic profile, i.e. a much shorter residence time in the blood compared to classical antibodies.[54] Furthermore, it is assumed that the small molecular weight of the vNAR domain implicates the opportunity to target epitopes otherwise only accessible to small molecules.[51]

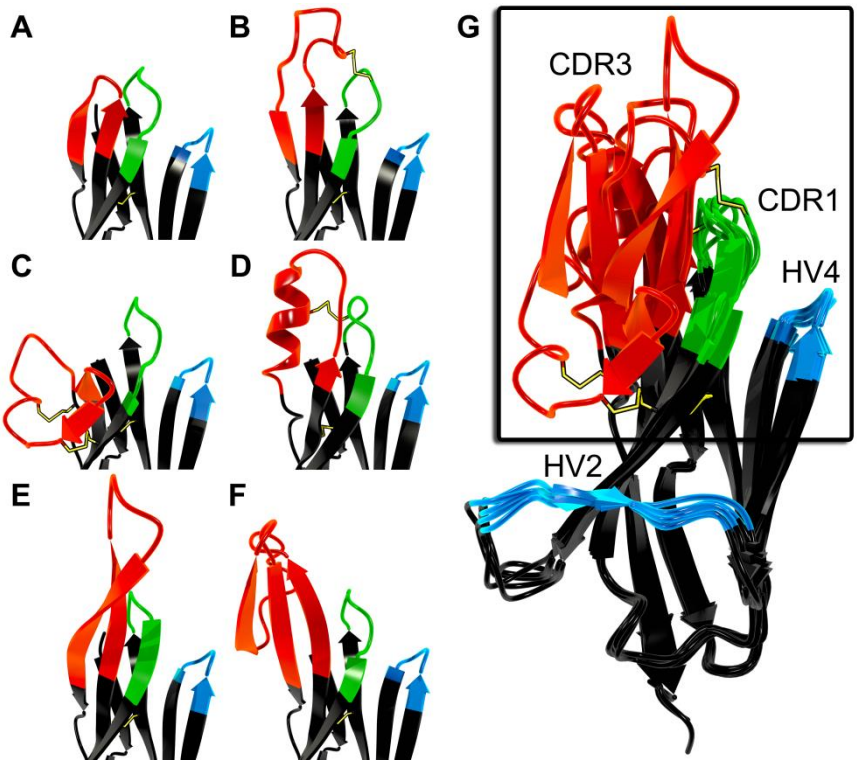


Figure 6. Examples of CDR3 variability in vNAR domains depicted in ribbon representation. (A) Short loop (type IV, pdb entry 4HGK).[103] (B) Large loop with one disulfide constraint (type II, pdb entry 2COQ).[25] (C) Highly constrained loop tethered by two cystine motifs (type I, pdb entry 1SQ2).[24] (D) Extended CDR3 forming an α- helical motif (type II, pdb entry 2I25).[99] (E) Extended CDR3 forming a two-stranded β-sheet (type IV, pdb entry 2Z8V).[127] (F) Extended CDR3 incorporating an amyloid-β p3 fragment (type IV, pdb entry 3MOQ).[129] (G) Overlay of structures A-F. Disulfide bonds are shown in yellow. Picture rendered with POV-Ray (www.povray.org/).

Table 1. vNAR domains selected against disease related targets. Adapted from ref. [105]

Target	Potential Application	K_D [nM]	Reference
VHSV	Anti-viral	-	[116]
TNF-α	Endotoxic shock	-	[130]
HSA	Half-life extension	0.66	[101]
HBeAg of HBV	Anti-viral	53 - 106	[124]
Ebola virus	Immunodiagnostic	-	[121]
Cholera toxin	Biosensor	83	[104]
Staphylococcal enterotoxin B	Sensor	10.2 - 107	[114]
Ricin	Sensor	299 ± 49	[114]
Botulinum toxin	Sensor	93 - 390	[114]
AMA1	Malaria diagnosis	15.8 - 358	[117,125]

VHSV: Viral hemorrhagic septicaemia virus, TNF: Tumor necrosis factor, HBeAg: Hepatitis B e antigen, HBV: Hepatitis B virus, HSA: Human serum albumin, AMA1: Malarial apical membrane antigen-1.

2.5.6 Re-formatting of vNAR domains

The simple single chain molecular architecture of vNAR domains affords the benefit of multiple re-formatting opportunities to tailor the final product for purpose. Many formats have been successfully proven, including monomeric, dimeric and trimeric (binding more than one target) in addition to Fc-based formats, with all demonstrating the inherent flexibility of these domains.[101,103,105,131]

The small size of vNAR domains leads to rapid renal clearance *in vivo* and represents a major drawback for non-imaging applications such as tumor-targeting. In contrast to this, a small size might be advantageous with regard to tumor penetration and ultimately the right format for the individual application needs to be determined. Fast glomerular filtration can be circumvented by multimerization of single vNAR domains, as shown by Müller *et al.*[101] Their investigations covered the *N*- as well as *C*-terminal fusion of a naïve vNAR domain to an anti-human serum albumin (HSA) vNAR originating from an immunized shark, isolated *via* phage display.[132] The fusion constructs retained high-affinity binding to HSA and exhibited significantly increased *in vivo* half-lifes compared to their unconjugated parental domains. The size of such dimeric formats is in the range of 25 kDa and compared to other antibody fragment formats such as scFvs, achieves double binding site capacity whilst retaining

increased affinities towards demanding, cryptic epitopes, which is a hallmark of vNAR proteins. In addition to dimeric fusions, a trimeric construct comprising the naïve vNAR domain at both termini of the anti-HSA vNAR displayed improved pharmacokinetics in *in vivo* studies in different species.[101] In another study conducted by Nuttall and colleagues, several approaches for the generation of bivalent shark antibodies with enhanced functional affinity were investigated.[131] Best affinities were obtained through *C*-terminal covalent or domain-mediated linkages.

As has been shown extensively for rodent mAbs, there is a plethora of rational as well as empirical humanization strategies available to reduce immunogenic responses caused by animal-derived immunoglobulins.[133-137] Rational design and the grafting of CDR loops of a xenogenic antibody onto a suitable human scaffold exhibiting a similar sequence has culminated in the development of several blockbuster pharmaceuticals routinely used in the clinic (e.g. trastuzumab, bevacizumab).[138]

Sequence identity of the IgNAR V domain with mammalian VH regions falls as low as 25 %.[100] In order to minimize the immunogenic potential of vNAR domains, Kovalenko and co-workers were able to engineer the aforementioned anti-HSA shark vNAR domain[101] by converting more than half of the framework amino acids to those of the human germline VK1 sequence DPK9.[103] This sequence bears the highest structural resemblance to the corresponding vNAR domain, and concomitantly represents one of the most stable human frameworks for downstream development. Determination of the binding constants of humanized vNAR variants yielded antigen affinities similar to those of the parental construct (14.8 nM *vs* 13.6 nM for the parental molecule). The opposite approach, to start with a structurally related human VH domain and convert that to a vNAR-like domain by clipping the CDR2 region, extending CDR3 and introducing stabilizing residues and additional disulfides may also be a viable alternative as already shown for the camelization of human VH domains.[139] To enhance the expression of humanized vNARs in mammalian cells they were *C*-terminally fused to human Fc domains.[103] This conjugation strategy can also contribute to the formation of dimers due to the interactions of two vNAR-Fc conjugates at the respective human constant domains. Fusions to human Fc, besides increasing the overall molecular weight and thus counteracting rapid renal clearance, can elicit *in vivo* immune effector functions and ultimately intensify the immune response *via antibody-dependent cellular cytotoxicity* (ADCC) and *complement-dependent cytotoxicity* (CDC). According to Kovaleva *et al.*, the humanized IgNAR V domain variant showed negligible immunogenicity in dendritic cell assays.[105] Notably, also no significant immunogenic effects were observed after subcutaneous application of the parental, non-humanized anti-HSA vNAR in rodents and non-human primates.[101,105] However, it remains to be scrutinized more meticulously how these proteins will behave when administered to patients in the scope of clinical trials.

Besides diagnostic and therapeutic applications, vNAR domains, due to their small size, high stability and their ability to sustain repeated cycles of unfolding and folding are also promising biomolecules for biotechnological applications to serve for example as high-affinity capturing agents for purification

of biomolecules or as tools for diagnostic applications. It was recently shown that vNAR fragments can be coupled covalently and site-specifically onto crystalline nanocellulose that serves as a protein-capturing nanoscaffold.[140] This evidence, coupled with the demonstrated stability and flexibility of vNAR domains, would predict more biotechnological applications can be expected to show up in the next years. Additionally, the vNAR domain can be utilized to gain information about pathological processes that at present are not completely understood. This was exemplified by Nuttall and colleagues, who grafted parts of the Aβ-peptide involved in Alzheimer's disease into CDR3 of a vNAR domain and thus solved the crystal structure of the amyloid-β p3 fragment (**Fig. 5F**).[129]

2.6 Antibody engineering and selection

Since the development of hybridoma technology in the 1970s by Köhler and Milstein[141], which is still state of the art in antibody technology, antibody development and engineering of therapeutic antibodies as well as of proteins in general has made continuous progress.[33,142] Nowadays, engineering and selection of completely human antibodies is amenable through the generation of transgenic animals bearing the genetic information for human antibody repertoires in the absence of mouse antibody genes.[143] Favored variants can be selected using the conventional hybridoma technology as well as *via* different display methodologies. Display technologies enable the identification of the respective candidates from libraries by the linkage between the protein variant to the coding information.[142] Thus, protein engineering and selection is not only limited to antibody molecules and it has been shown that an unprecedented range of proteins and enzymes can be manipulated as desired.[69,142,144-148]

In general there are five main display technologies, referred to as phage display, cellular display systems, ribosomal display, mRNA display and DNA display.[142] Compared to the other aforementioned display systems, cellular screening harbors the advantages of single-cell on-line and real-time analysis as well as characterization of individual library members. Moreover, using eukaryotic screening platforms, displayed protein variants are believed to be folded correctly, owing to the fact that every library member has to pass intrinsic quality control machineries for proper folding before getting displayed on the surface of the host cell.[149] Different cellular display platforms, for instance bacterial display, yeast surface display and various mammalian display systems have recently been reviewed by Kolmar and co-workers.[149] Within these studies, yeast surface display (YSD) was used as platform technology and will be scrutinized in more detail.

Yeast surface display was introduced by Boder and Wittrup in 1997.[150] In this original approach, the Aga1p mating adhesion receptor is integrated into the genome of *Saccharomyces cerevisiae* under the control of a galactose-inducible promoter. The gene of interest is cloned as a fusion to the episomal encoded AGA2. Both subunits are approximately expressed to the same extent and associate

covalently *via* two disulfide bonds following galactose-induction in the endoplasmatic reticulum. Due to a *C*-terminal glycosylphosphatidylinositol anchor (GPI attachment signal) of Aga1p, the fusion-construct containing the protein of interest, is displayed on the cell wall in a copy number of approximately 10^4 – 10^5 respective protein variants (**Fig. 7**).[149,150] The protein of interest i.e. library candidate can be flanked by several epitopes for the detection of surface expression. Besides the Aga1p – Aga2p system, various other cell wall proteins have proven to be successful for yeast surface display.[151,152] It needs to be mentioned that aside from the platform technology, pioneered by Boder and Wittrup, several other display systems are routinely used to engineer proteins, including display technologies using *Pichia pastoris* as host.[149,153,154] Yeast surface display is compatible with fluorescence-activated cell sorting (FACS) which enables control over the selection process, i.e. the above mentioned on-line and real-time analysis. This allows to fine-tune and to discriminate very sensitively for affinity and stability. Another benefit of yeast surface display is the correlation between display level on the yeast surface and thermal stability as well as soluble expression level, enabling the rapid isolation of stable proteins with desired properties such as affinity.[155] Additionally, affinities and stability of individual library candidates can be determined on the yeast surface, which is a convenient way to characterize a large number of candidates without the need for soluble expression.[146,156] Finally, antibody libraries are suggested to be sampled more comprehensively with yeast surface display than with phage display.[157]

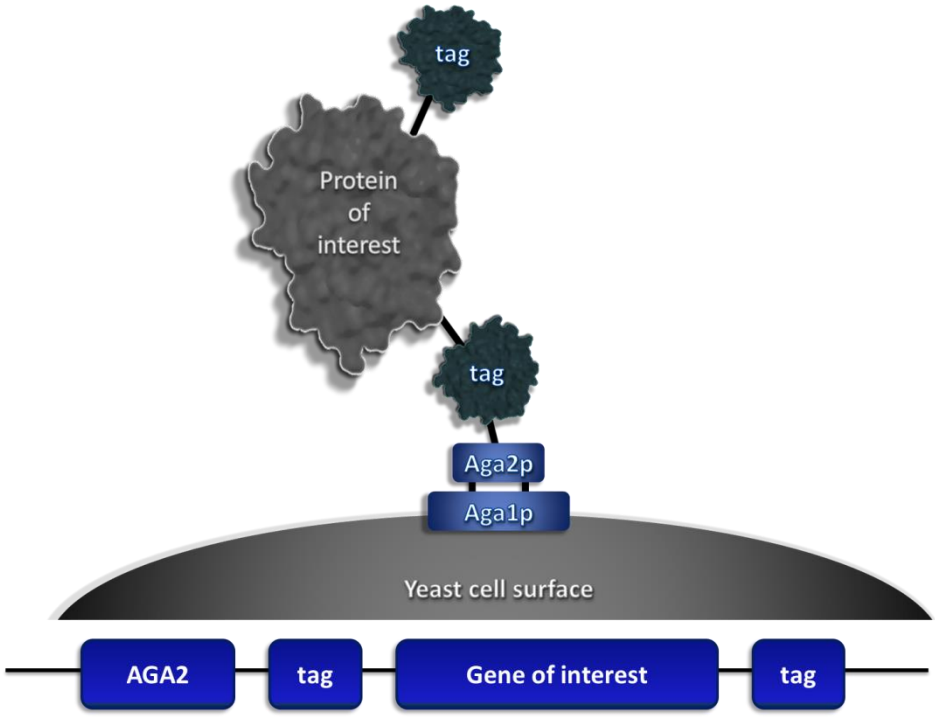


Figure 7. Simplified illustration of yeast surface display according to Boder and Wittrup.[150] The protein of interest is fused to the *C*-terminus of Aga2p. Below, the episomal genetic arrangement is shown. Tag: Epitope for the detection of surface presentation *via* immunofluorescence staining.

2.7 Target proteins

In the following section the antigens used in this study are briefly described. Although antigens were used as proof-of-concept target proteins, most - if not all - are of therapeutic and/or diagnostic relevance.

2.7.1 Epithelial cell adhesion molecule (EpCAM)

Epithelial cell adhesion molecule EpCAM is expressed on a subset of epithelia where it mediates epithelial-specific intercellular cell-adhesion *via* oligomerization.[158,159] It also plays an important role to the formation of tight junctions.[160,161] Though, it was also shown that EpCAM can have a negative impact on E-cadherin-mediated cell-cell adhesion, thus promoting cell migration and motility.[159,162] EpCAM is overexpressed on malignant cells in a variety of different tumor entities. Furthermore, it was identified as a marker for cancer initiating cells.[163,164] In certain tumor types, overexpression of EpCAM is correlated with an advanced stage of cancer and with worse overall survival as well as risk for recurrence.[158,163] EpCAM is a type I transmembrane glycoprotein and consists of a large extracellular domain (EpEX, 27 kDa), a transmembrane domain and a short intracellular domain (EpICD).[163] Recently, Lenarcic and co-workers were able to solve the crystal structure of EpEX.[159] The group demonstrated that the extracellular domain of EpCAM forms a dimer on the cell-surface and provided a model in which intercellular cell-cell contacts are formed through EpCAM tetramerization mediated by membrane-distal regions.[159]

2.7.2 Receptor tyrosine kinase EphA2

Receptor tyrosine kinases are key players in cell signaling pathways in carcinogenesis.[165,166] The Eph family, including EphA2, represents the largest group among receptor tyrosine kinases.[167] Physiological roles of Eph members include participation in the development of neuronal networks, axon guidance, formation and remodeling of synaptic connections and nervous system repair.[168] However, EphA2 is overexpressed and functionally altered in a multitude of different cancers, including breast cancer, melanomas, ovarian cancer, lung cancer, gliomas, urinary bladder cancer, prostate cancer as well as esophageal, renal, colon and vulvar cancers.[167,169,170] Besides being a marker for metastatic behavior, it is known that EphA2 overexpression directly promotes an aggressive phenotype and its overexpression can induce malignant transformation.[169] Consistently, high EphA2 expression in tumors is correlated with poor prognosis and recurrence owing to its metastatic potential.[167] EphA2 is a large molecule of approximately 130 kDa and consists of an extracellular conserved *N*-terminal ligand binding domain (EphA2 interacts with any of eight different ephrinA-family ligands [169]), followed by a cysteine-rich domain as well as an EGF-like motif and two

fibronectin type III repeats. The extracellular part is followed by a membrane-spanning region and the cytoplasmic region that contains a juxtamembrane region, a tyrosine kinase domain, a sterile alpha motif and a post synaptic domain.[167]

2.7.3 Interleukin-8 (CXCL8)

Chemokines are a family of structurally related small proteins (8-10 kDa) involved in the regulation of immune processes. CXCL8 is a pro-inflammatory chemokine which acts on leukocytes and endothelial cells. Interleukin-8 (IL-8) as well as other pro-inflammatory cytokines have been associated with chronic inflammation and inflammatory diseases.[171] IL-8 interacts with receptors CXCR1 and CXCR2 with high affinities.[172] Besides its role in inflammation,[173] there is evidence that IL-8 functions in the tumor microenvironment.[172] The cytokine responds to the activation of NF-κB, which is commonly activated in cancer cells. IL-8 is believed to stimulate angiogenesis, cell proliferation, survival as well as migration and metastasis.[174] Additionally, the cytokine mediates tumor cells to undergo epithelial-mesenchymal transition. Thus, IL-8 and its receptors emerged as promising targets for cancer therapy.[172,175]

2.7.4 Cluster of differentiation 3 (CD3) subunit CD3ε

Cluster of differentiation 3 is a T-cell co-receptor composed of four different chains, CD3γ, CD3δ, CD3ζ and CD3ε. The multi-protein complex associates non-covalently with the T-cell receptor, representing the T-cell receptor complex.[176] This complex couples antigen-recognition, mediated by the variable domains of the T-cell receptor with intracellular signaling and T-cell activation, mainly through the ζ-chain.[177] CD3, especially subunit CD3ε is now recognized as a promising target for T-cell activation for bi-specific antibodies.[82,178] bsAbs for this kind of immunotherapy bind with one arm a tumor-associated antigen, while with the other antigen-binding site the T-cell is engaged through CD3ε. Thus, the tumor gets cross-linked to T-cells and the latter are getting simultaneously activated for their cytotoxic activity.[178] This cytotoxic activity is owing to the formation of the cytolytic synapse i.e. the release of perforin and granzymes by the cytotoxic T-cell and their serial target cell lysis capacity.[179,180] To this end, this prospective approach of T-cell engagement already reached approval[81] and several molecules are currently assessed in clinical trials.[179]

2.7.5 Human Fcγ

Fcγ is the Fc-part of IgG, the most abundant isotype found in the serum of humans. Due to its long serum half-life [31], binding to IgG could significantly increase circulation time of small proteins, such as shark vNAR antibody domains. As aforementioned, in a similar approach, executed by Barelle and colleagues, binding of a vNAR domain to human serum albumin (HSA) drastically increased the serum half-life of this molecule, partially due to HSA recycling, mediated by FcRn.[101]

2.8 Aim of the work

This work aimed at generating high-affinity antigen binding vNAR molecules from the bamboo shark (*Chiloscyllium plagiosum*) against different disease-related antigens. Thus far, the bamboo shark has not been described in literature for the selection of target-specific vNAR molecules. Hence, the main objective was to establish a platform for antigen-specific IgNAR V domain generation by resembling the natural immune response in sharks using yeast surface display as technology for antibody selection.

For this, binders against the aforementioned target-proteins needed to be selected from a CDR3-diversified repertoire. Affinity maturation should be demonstrated by the diversification of CDR1 of target-enriched antigen binding vNAR populations and sublibrary screening with enhanced stringency, to obtain binders with optimized affinities. This strategy kind of mimics the natural mechanism of affinity maturation by somatic hypermutation in sharks.

Additionally, we also set out to generate a new antigen-binding site into the vNAR scaffold that acts autonomously from the conventional paratope. Taking a closer look at hypervariable loop 2 (HV2), previous studies showed that HV2 is prone to mutations during somatic hypermutation. However, we perceived that this loop is located distantly from the conventional paratope, composed of CDR3, CDR1 as well as hypervariable loop 4 (**Fig. 8**). To the best of our knowledge, thus far evidence is missing that HV2 ever contributed to antigen binding. Essentially, we attempted to utilize this loop (HV2) to generate a new antigen-binding site, resulting in a bi-specific vNAR with the capability to address two distinct targets. Ultimately, a bi-specific vNAR molecule would be the smallest bi-specific antibody domain known in the animal kingdom.

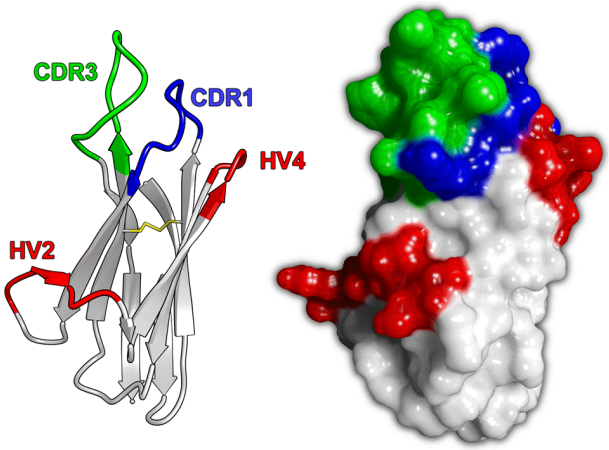


Figure 8. Structure of a vNAR binding domain (left: ribbon model, right: surface) indicating variable regions CDR1, HV2, HV4, and CDR3. Disulfide bond is shown in yellow. The model was generated on the basis of a published X-ray structure (PDB-ID: 1VES) using YASARA structure [107] and rendered with POVRay using radiosity (left) and subsurface light transport (right). Figure was constructed in close cooperation with Martin Empting.

3. MATERIAL

3.1 Bacterial strains, yeast strains and cell lines

Bacterial strains – *Escherichia coli*

BMH 71-18 - *thi, supE, Δ(lac-proAB), [mutS::Tn10], [F', proAB, laqI q ZΔM15]*

DH5α - *F- f80lacZΔM15 Δ(lacZYA-argF)U169 recA1 endA1 hsdR17(rk-, mk+) phoAsupE44 thi-1 gyrA96 relA1 λ-*

TOP 10 - F- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Δ80*lacZ* ΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*araleu*) 7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G

Yeast strains – *Saccharomyces cerevisiae*

EBY100 - *a GAL1-AGA1::URA3 ura3-52 trp1 leu2Δ200 his3Δ200 pep4::HIS2 prb11.6R can1 GAL, Trp-Leu-*

Cell lines

CHO	Chinese hamster ovary Cultivation at 37°C, 5 % CO_2
EBC-1	Human lung squamous cell carcinoma Cultivation at 37°C, 5 % CO_2
HEK-293	Human embryonic kidney Cultivation at 37°C, 5 % CO_2
HT-29	Human colorectal adenocarcinoma Cultivation at 37°C, 5 % CO_2
T47D	Human mammary gland, breast tumor Cultivation at 37°C, 5 % CO_2

3.2 Plasmids

3.2.1 pCT

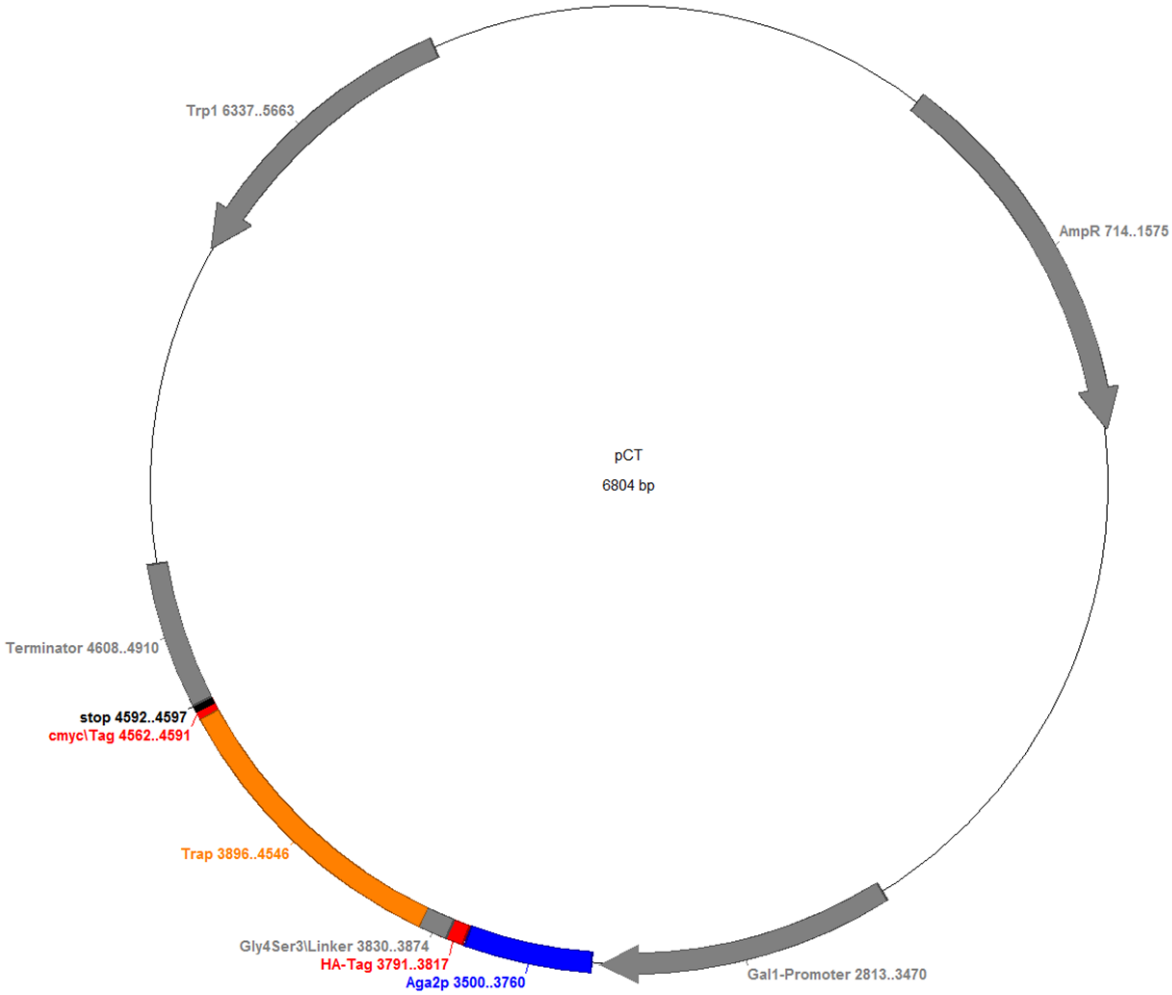


Figure 9. Vector map for pCT plasmid. Arrows indicate orientations of genetic elements. Trp1: gene sequence for phosphoribosyl anthranilate isomerase (auxotrophic marker). AmpR: gene for β-lactamase. Gal1-Promoter: Gal1/10 promoter region. Aga2p: coding sequence for AGA2 protein. HA-Tag: HA epitope. Gly4Ser3/Linker: glycine-serine linker sequence. Trap: partial DNA sequence of TRAP protein (not in frame with Aga2p). cmyc/tag: cMYC epitope coding DNA sequence. stop: tandem stop codon. Terminator: terminator sequence.

3.2.2 pMX-EETI

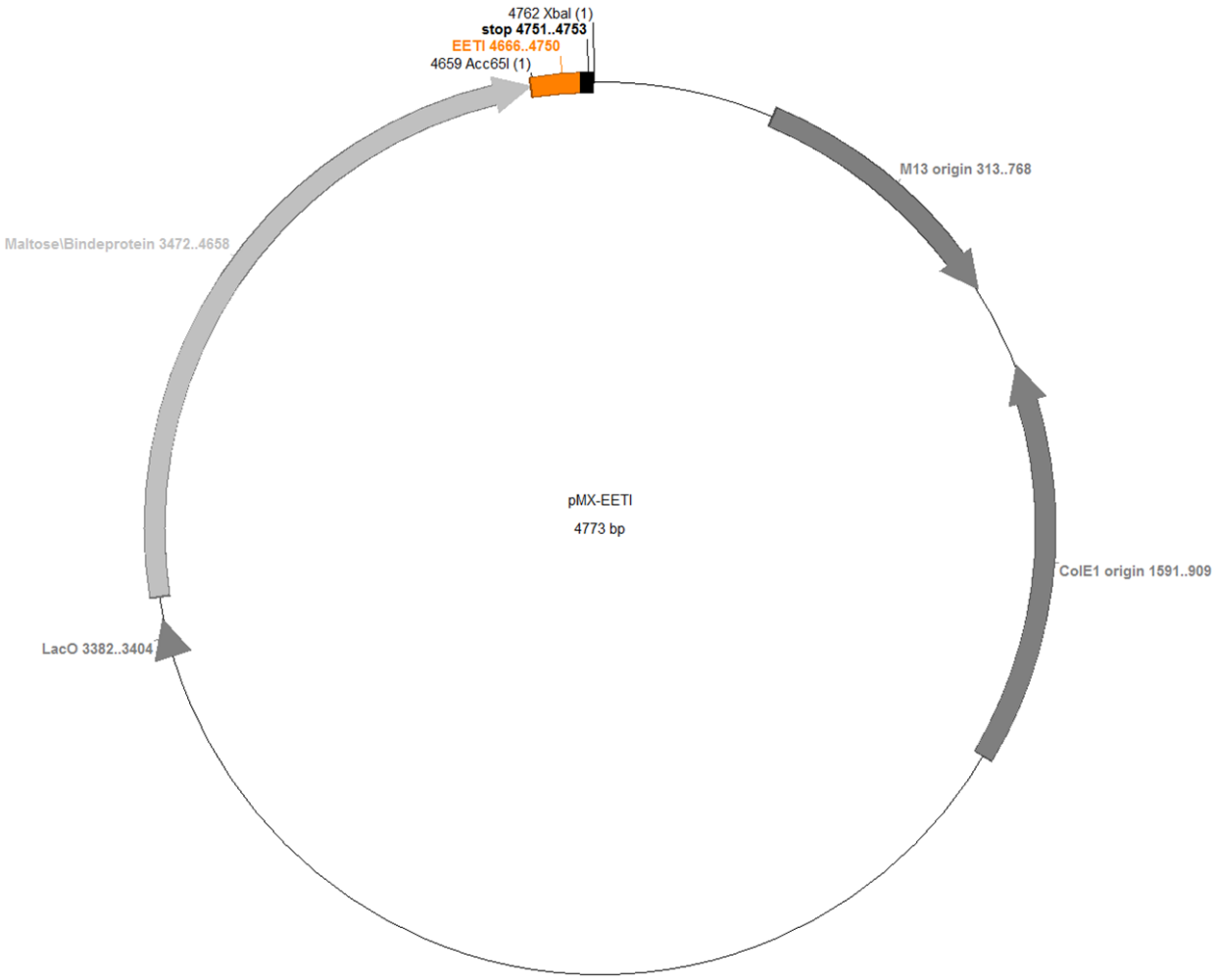


Figure 10. Vector map for pMX-EETI plasmid. Arrows indicate orientations of genetic elements. LacO: gene sequence for lac operon. Maltose/Bindeprotein: gene for MalE (which is expressed as fusion with protein of interest). EETI: coding sequence for EETI protein. stop: stop codon. Acc65I: recognition sequence for restriction enzyme *Acc* 65I. XbaI: recognition sequence for restriction enzyme *Xba* I. M13 origin: M13 bacteriophage origin of replication. ColE1 origin: bacterial origin of replication.

3.2.3 pMX-vNAR

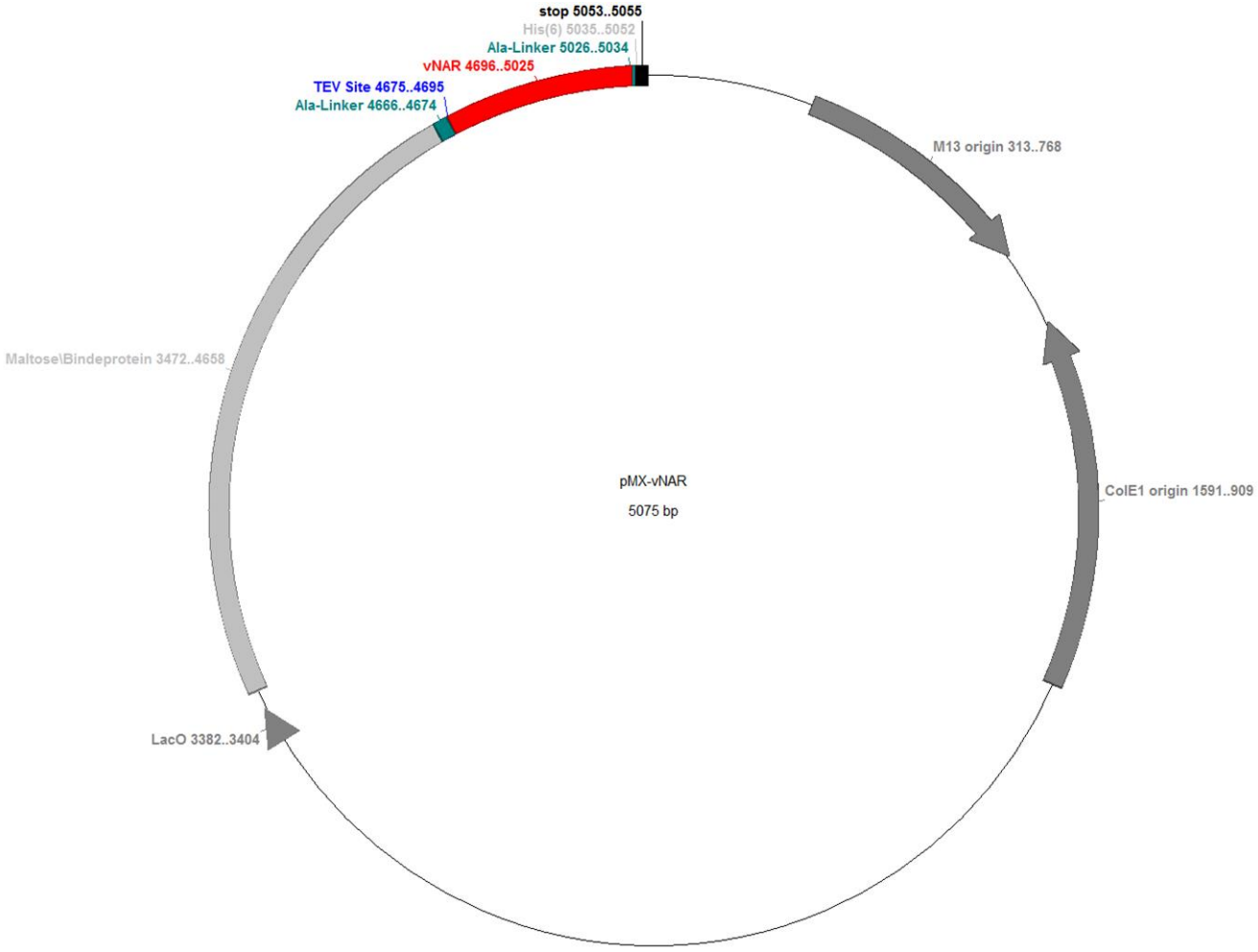


Figure 11. Vector map for pMX-vNAR plasmid. Arrows indicate orientations of genetic elements. LacO: gene sequence for lac operon. Maltose/Bindeprotein: gene for MalE (which is expressed as fusion with protein of interest). Ala-Linker: sequence for triple alanine. TEV Site: recognition sequence for tabacco etch virus protease cleavage. vNAR: sequence for *E. coli* codon-optimized vNAR sequence of α-EpCAM-vNAR_5005 stop: stop codon. Acc65I: recognition sequence for restriction enzyme *Acc* 65I. XbaI: recognition sequence for restriction enzyme *Xba* I. M13 origin: M13 bacteriophage origin of replication. ColE1 origin: bacterial origin of replication.

3.2.4 pExpress-Fc

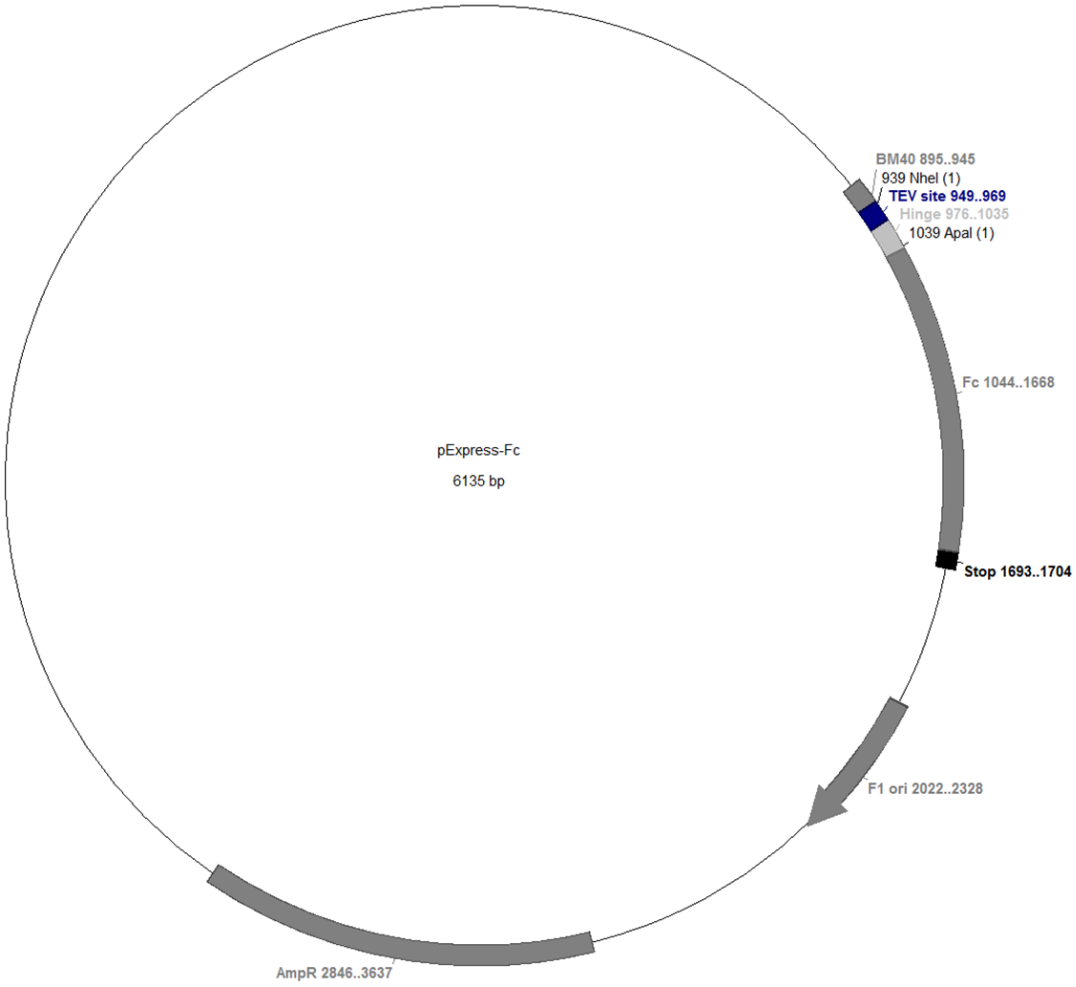


Figure 12. Vector map for pExpress-Fc plasmid. Arrows indicate orientations of genetic elements. F1 ori: phage-derived origin of replication. AmpR: gene for β-lactamase. BM40: signal peptide sequence. NheI: recognition sequence for restriction enzyme *Nhe* I. TEV site: : recognition sequence for tabacco etch virus protease cleavage. Hinge: coding sequence for hinge region. ApaI: recognition sequence for restriction enzyme *Apa* I. Fc: coding sequence for domains CH2 and CH3 of human IgG (non-glycosylated, Asn297 replaced by Ala297). stop: triple stop codon. Promoter sequence of cytomegalovirus not shown.

3.3 Enzymes, proteins, nucleic acids and protein ladder

2-Log DNA ladder	New England Biolabs (NEB), Beverly, USA
Acc 65I	New England Biolabs (NEB), Beverly, USA
Anti-CD326 (EpCAM)-PE	MACS Miltenyi Biotec, Bergisch-Gladbach, Germany
Anti-cMyc- antibody	Made in-house
Anti-HA-antibody	eBioscience, San Diego, USA
Anti-Hu IgG (Fcγ-specific) PE	eBioscience, San Diego, USA
Anti-human-Fab-488 conjugate	Jackson Immuno Research, Suffolk, UK
Anti-Mouse IgG Biotin	Sigma Aldrich, Steinheim, Germany
Anti-Mouse IgG FITC	Sigma Aldrich, Steinheim, Germany
Apa I	New England Biolabs (NEB), Beverly, USA
Bam HI	GERBU Biotechnik, Heidelberg, Germany
Bovine serum albumin (BSA)	GERBU Biotechnik, Heidelberg, Germany
dNTPs	Carl Roth GmbH, Karlsruhe, Germany
HTRA1	Made in-house
Henn egg white lysozyme	MP Biomedicals, LCC. Illkirch, France
(mm) EGFR	Merck (EMD) Serono, Darmstadt, Germany
(mm) EpCAM	ACRO Biosystems, Newark, USA
(mm) EphA2	Sino Biological Inc., Beijing, USA
Nhe I	New England Biolabs (NEB), Beverly, USA
Oligonucleotide primers	Sigma Aldrich, Steinheim, Germany
Penta-His Alexa Fluor 488 conjugate	Qiagen, Hilden, Germany
Prestained Protein Marker, Broad Range	New England Biolabs (NEB), Beverly, USA
Reverse transcriptase (Omniscript RT)	Qiagen, Hilden, Germany
(rh)CD3ε	ACRO Biosystems, Newark, USA
(rh) EpCAM	ACRO Biosystems, Newark, USA
(rh) EphA2	R&D Systems, Minneapolis, USA
(rh) Fcγ (glycosylated)	Made in-house
(rh) Fcγ (non-glycosylated)	Made in-house
(rh) IL-8	Prof. Schmitz, TU Darmstadt, Germany

Streptavidin APC (SAPC)	eBioscience, San Diego, USA
Streptavidin-R-phycoerythrin conjugate (SPE)	Sigma-Aldrich, Steinheim, Germany
T4 DNA ligase	New England Biolabs (NEB), Beverly, USA
Taq Polymerase	New England Biolabs (NEB), Beverly, USA
TEV protease	Made in-house
tGFP	Made in-house
Trypsin-EDTA	Gibco(Invitrogen), Karlsruhe, Germany
Xba I	New England Biolabs (NEB), Beverly, USA

3.4 Oligonucleotides

Name	**Sequence (5′–3′)**
bamboo/nat_lo	WTTCACAGTCASARKGGTSCC
bamboo/nat_up	ATGGCCSMACGGSTTGAACAAACACC
CDR1rand_up	ACCATCAATTGCGTCCTAAAA$(X)_5$TTGGGTAGCACGTACTGGTATTTCACAAAGAAG
CDR3rand/Fr4_lo	WTTCACAGTCASARKGGTSCCSCCNCCTTCAAT$(X)_{12}$CGCTTCACAGTGATATGTACC
FR1_up	ATGGCCGCACGGCTTGAACAAACACCGACAACGACAACAAAGGAGGCAGGCGAATCACTGACCATCAATTGCGTCCTAA
FR1/CDR1/Tyr_up	ACCATCAATTGCGTCCTAAAAGGTTCCRNMTATGBATTGGGTANMACGTACTGGT
FR3_lo	CGCTTCACAGTGATATGTACC
GR_up	GTGGTGGTGGTTCTGCTAGCATGGCCGCACGGCTTGAACA
GR_lo	ATAAGCTTTTGTTCGGATCCWTTCACAGTCASARKGGTSCCSCCNCC
HV2_B1mut_lo	TGTGTCCGAGTAGCGTCCGCCCATCCATTT

HV2_B1mut_up	AAATGGATGGGCGGACGCTACTCGGACACA
HV2_SOE_lo	GCCCTTCTTTGTGAAATACCA
HV2_SOE_rand_up	TGGTATTTCACAAAGAAGGGC$(X)_9$GCGGACGATACTCGGACACA
pCT_Seq_lo	GCGCGCTAACGGAACGAAAAATAGAAA
pCT_Seq_up	GCGGCGGTTCCAGACTACGCTCTGCAGGCT
pEXPR_vNARB1mut_loI	GGGTCTTGTCGCAGCTCTTGGGCTCGCTTCCGCTCTGGAAGTACAGGTT CTCTTTCACAGTCACAGTGGTC
pEXPR_vNARB1mut_loII	GAACACGCTGGGCCCGCCCAGCAGTTCAGGGGCAGGGCAGGGAGGAC AGGTGTGGGTCTTGTCGCAGCTCT
pEXPR_vNAR_loI	GCAGGGAGGACAGGTGTGGGTCTTGTCGCAGCTCTTGGGCTCGCTTCC ATTCACAGTCACAAGGGTCCC
pEXPR_vNAR_loII	GGAACACGCTGGGCCCGCCCAGCAGTTCAGGGGCAGGGCAGGGAGGA CAGGTGTGGGTCTTGTCGCA
pEXPR_vNAR_up	TGGCCGGAAGGGCGCTAGCCGCTGAGAACCTGTACTTCCAGAGCATGG CCGCACGGCTTGAACA
pMX_vNAR_lo	TCGACCTCTAGAGCGGCCGCTTAATGATGATGATGATGATGCGCCGCC GCATTCACAGTCACAAGGGTCCC
pMX_vNAR_up	AATTCGAGCTCGGTACCCGCGGCGGCGGAAAACCTGTATTTTCAGAGC ATGGCCGCACGGCTTGAACAA

X: triplet codon for all natural amino acids w/o Cys; B: G/C/T; K: G/T; M: A/C; N: A/G/C/T; R: A/G; S: G/C; W: A/T.

3.5 Chemicals

2-Mercaptoethanol	Carl Roth GmbH, Karlsruhe, Germany
Acetic Acid	Carl Roth GmbH, Karlsruhe, Germany
Acrylamide/Bisacrylamide (37,5:1)	Carl Roth GmbH, Karlsruhe, Germany

Agar-Agar	Carl Roth GmbH, Karlsruhe, Germany
Agarose Ultra-Quality	Carl Roth GmbH, Karlsruhe, Germany
Ammonium acetate	Carl Roth GmbH, Karlsruhe, Germany
Ammonium hydrogene carbonate	Carl Roth GmbH, Karlsruhe, Germany
Ampicillin, sodiumsalt	Carl Roth GmbH, Karlsruhe, Germany
APS	Carl Roth GmbH, Karlsruhe, Germany
Bacillol plus	Carl Roth GmbH, Karlsruhe, Germany
Bromophenol blue	Carl Roth GmbH, Karlsruhe, Germany
Calcium chloride	Carl Roth GmbH, Karlsruhe, Germany
Casaminoacids	Becton Dickinson GmbH, Heidelberg, Germany
Citric acid	Carl Roth GmbH, Karlsruhe, Germany
Chloramphenicol	Carl Roth GmbH, Karlsruhe, Germany
Chloroform	Fisher Scientific, Schwerte, Germany
Coomassie brilliant blue R250	Carl Roth GmbH, Karlsruhe, Germany
Di-sodium hydrogen phosphate	Carl Roth GmbH, Karlsruhe, Germany
Dithiothreitol	Carl Roth GmbH, Karlsruhe, Germany
DMSO	Sigma Aldrich, Steinheim, Germany
Ethanol	Carl Roth GmbH, Karlsruhe, Germany
EDTA	Applichem GmbH, Darmstadt, Germany
Ethidium bromide	Carl Roth GmbH, Karlsruhe, Germany
Fetal bovine serum	Sigma-Aldrich, Steinheim, Germany
Galactose	Sigma-Aldrich, Steinheim, Germany
Gel loading dye (6x)	New England Biolabs (NEB), Beverly, USA
Glucose	Carl Roth GmbH, Karlsruhe, Germany
Glycerol	Carl Roth GmbH, Karlsruhe, Germany
Glycerin	Carl Roth GmbH, Karlsruhe, Germany
HCl	Carl Roth GmbH, Karlsruhe, Germany
HDgreen Plus DNA Stain	INTAS, Göttingen, Germany
Imidazole	Carl Roth GmbH, Karlsruhe, Germany
Isopropyl alcohol	Carl Roth GmbH, Karlsruhe, Germany
IPTG	Carl Roth GmbH, Karlsruhe, Germany
Lithium acetate	Carl Roth GmbH, Karlsruhe, Germany

L-Glutamine	Sigma Aldrich, Steinheim, Germany
Magnesium chloride	Carl Roth GmbH, Karlsruhe, Germany
Methanol	Carl Roth GmbH, Karlsruhe, Germany
Nickel chloride	Merck KGaA, Darmstadt, Germany
Paraffin oil	Carl Roth GmbH, Karlsruhe, Germany
Penicillin/Streptomycin (100x)	Life Technologies, Darmstadt, Germany
Phenol	Carl Roth GmbH, Karlsruhe, Germany
Peptone/Tryptone	Carl Roth GmbH, Karlsruhe, Germany
Polyethylene glycol 8000	Carl Roth GmbH, Karlsruhe, Germany
Potassium chloride	Carl Roth GmbH, Karlsruhe, Germany
Potassium di-hydrogen phosphate	Carl Roth GmbH, Karlsruhe, Germany
Potassium hydrogen phosphate	Carl Roth GmbH, Karlsruhe, Germany
Potassium hydroxide	Carl Roth GmbH, Karlsruhe, Germany
Sucrose	Carl Roth GmbH, Karlsruhe, Germany
SDS, ultra pure	Carl Roth GmbH, Karlsruhe, Germany
Sodium chloride	Carl Roth GmbH, Karlsruhe, Germany
Sodium hydroxide	Carl Roth GmbH, Karlsruhe, Germany
Sodium di-hydrogen phosphate di-hydrate	Carl Roth GmbH, Karlsruhe, Germany
Sorbitol	Carl Roth GmbH, Karlsruhe, Germany
Sulfo-NHS-LC-Biotin	Thermo Fisher Scientific, Geel, Belgium
Sypro Orange Gel Protein Stain	Sigma Aldrich, Steinheim, Germany
TEMED	Merck KGaA, Darmstadt, Germany
TRIS	Carl Roth GmbH, Karlsruhe, Germany
Tri-sodium citrate	Carl Roth GmbH, Karlsruhe, Germany
Yeast extract	Carl Roth GmbH, Karlsruhe, Germany
Yeast Nitrogen Base	Becton Dickinson GmbH, Heidelberg, Germany
TWEEN-20	Carl Roth GmbH, Karlsruhe, Germany
Triton X-100	Carl Roth GmbH, Karlsruhe, Germany

3.6 Solutions and buffers

1x SDS running buffer	Tris 50 mM, Glycin 190 mM, SDS 1 g/l
4x PAGE buffer A	Tris/HCl 3 M (pH 8.85), SDS 4 g/l
4x PAGE buffer B	Tris/HCl 0.5 M (pH 6.8), SDS 4 g/l
5x PAGE loading buffer	Tris/HCl 0.25 M (pH 8), SDS (7.5 % (w/v), Glycerin 25 % (w/v), bromphenol blue 0.25 mg/ml, 2-mercaptoethanol 12.5 % (w/v)
APS-stock	10 % (w/v)
Coomassie staining solution	Acetic acid 10 % (v/v), methanol 40 % (v/v), Coomassie brilliant blue R250 0.25% (w/v)
Electroporation buffer (yeast)	Sorbitol 1 M, calcium chloride 1 mM
IMAC EDTA solution	EDTA 100 mM
IMAC elution buffer	Imidazol 1M in PBS (pH 7.4)
IPTG stock	IPTG 1 M
Kinetics buffer (FortéBio)	PBS (pH 7.4), 0.1 % BSA (w/v), 0.02 % Tween-20
LiAc/DTT buffer	Lithium acetate 100 mM, Dithiothreitol 10 mM
Nickel chloride solution	Nickel chloride 100 mM
Periplasm buffer	Sucrose 0.5 M, Tris/HCl 0.1 M, EDTA 1 mM
Phosphate-buffered saline (PBS)	8.1 g/L NaCl, 0.75 g/L KCl, 1.13 g/L Na_2HPO_4 and 0.27 g/L KH_2PO_4, pH 7.4
Protein A chromatography running buffer	Sodium di-hydrogen phosphate di-hydrate 1.08 g/L, Di-sodium hydrogen phosphate 1.72 g/L (pH 7.0)
Protein A chromatography elution buffer	Citric acid 17.2 g/L, Tri-sodium citrate 5.28 g/L (pH 3.0)
Protein A chromatography collection buffer	Tris/HCl 1 M (pH 9.0)

TAE buffer (50x)	Tris 2 M, acetic acid 1M, EDTA 0.1 M
TE buffer	Tris/HCl 10 mM (pH 8.0)
TEV protease reaction buffer (10x)	500 mM Tris/HCl (pH 7.5), 1.5 M NaCl
Trypan blue solution	Trypan blue 4 % (w/v)
Yeast freezing solution	Glycerol 2 % (v/v), yeast nitrogen base 0.67 % (w/v)
Yeast lysis buffer	Tris/HCl 10 mM (pH 8), NaCl 100 mM, SDS 1 % (w/v), EDTA 1 mM, Triton X-100 2 % (v/v)

3.7 Cell culture media

E. coli

dYT	Yeast extract 10 g/L, peptone/tryptone 16 g/L, NaCl 5 g/L

S. cerevisiae

SD-CAA	Yeast nitrogen base without amino acids and ammonium sulfate 1.7 g/L, ammonium sulfate 5 g/L, 5 g/L casamino acids, 20 g/L dextrose, 8.6 g/L $NaH_2PO_4 \times H_2O$, 5.4 g/L Na_2HPO_4
SG-CAA	Yeast nitrogen base without amino acids and ammonium sulfate 1.7 g/L, ammonium sulfate 5 g/L, 5 g/L casamino acids, 20 g/L galactose, 8.6 g/L $NaH_2PO_4 \times H_2O$, 5.4 g/L Na_2HPO_4, polyethylene glycol 8000 10 % (w/v)
YPD	20 g/L tryptone, 20 g/L dextrose and 10 g/L yeast extract

human cell line HEK293

Freestyle™ 293 expression medium	Life Technologies, Darmstadt, Germany

3.8 Laboratory materials and kits

5 ml tubes for FACS	Saarstedt, Nümbrecht, Germany
50 ml centrifuge tubes with screw caps	Carl Roth GmbH, Karlsruhe, Germany
emPCR Kit *(Lib-A)*	454 Life Sciences (Roche), Branford, USA
Expi293™ Expression System Kit	Life Technologies, Darmstadt, Germany
Filtropure filter devices	Saarstedt, Nümbrecht, Germany
FortéBio tips	FortéBio, Menlo Park, California, USA
GS Junior Titanium PicoTiterPlate Kit	454 Life Sciences (Roche), Branford, USA
GS Junior Titanium Sequencing Kit	454 Life Sciences (Roche), Branford, USA
Dialysis membranes	Carl Roth GmbH, Karlsruhe, Germany
MicroBeads	Beckman Coulter GmbH, Krefeld, Germany
Mikrotiter plates 96-well (for FortéBio)	Greiner Bio-One GmbH, Frickenhausen, Germany
Omniscript® Reverse Transcription Kit	Qiagen, Hilden, Germany
Protein Desalting Spin Columns	Thermo Scientific, Rockford, USA
Protein purification columns, miscellaneous	GE Healthcare Europe GmbH, Freiburg, Germany
Sphero™ rainbow calibration particles	Spherotech Inc, Lake Forest, USA
Tri® Reagent BD	Sigma Aldrich, Steinheim, Germany
Wizard® Plus SV Midipreps DNA Purification System	Promega, Mannheim, Germany
Wizard® Plus SV Minipreps DNA Purification	Promega, Mannheim, Germany

System	
Wizard® SV Gel and PCR Clean-Up System	Promega, Mannheim, Germany

Further materials comprised common laboratory devices.

3.9 Equipment

Äkta Basic UV900 P900 Frac 900, Unicorn 3.1 Software	GE Healthcare Europe GmbH, Freiburg, Germany
BioRad 96CFX RT-PCR detection system	Bio-Rad, München, Germany
Fluorescence Activated Cell Sorter (FACS), MoFlo® Cytometer with Cyclone sorting unit, Summit® v4.3 Software package	Dako Cytomation, Fort Collins, USA
FortéBio octet red	FortéBio, Menlo Park, California, USA
Gel documentation system, Gel Jet Imager	INTAS Science Imaging Instruments GmbH, Göttingen
Gene Pulser® und pulse controller	BioRad, München
GS Junior System (454 Sequencing)	454 Life Sciences (Roche), Branford, USA
Microscope Axioskop	Carl Zeiss Microscopy GmbH, Jena, Germany
PCR Cycler Eppendorf Mastercycler	Eppendorf, Hamburg, Germany

Further equipment comprised common laboratory instrumentation.

3.10 Animals

Three specimen of the bamboo shark (*Chiloscyllium plagiosum*) in the age of approximately eight months were obtained from the Vivarium of the Staatliches Museum für Naturkunde, Karlsruhe, Germany. Procedures were conducted in accordance with the national laws § 4 Abs. 3 of the German Tierschutzgesetz (TierSchG, animal welfare act). Permission number: V 54 – 19 c 20 15 (1) Gl 18/19 Nr. A 35/2011, Regierungspräsidium Giessen, Germany (Regional council Giessen).

4. METHODS

4.1 Analysis of the natural repertoire of the bamboo shark

Blood samples of one individual of the bamboo shark (*C. plagiosum*) were harvested from the caudal vein and subsequently transferred into Tri® Reagent BD (Sigma Aldrich). Total RNA was extracted from whole blood according to the manufacturer's protocol (Sigma Aldrich). The Omniscript® Reverse Transcription Kit (Qiagen) was used for cDNA synthesis as described in the manufacturer's protocol. Five reactions of pooled total RNA from one individual were carried out in parallel, each containing about 2 µg of total RNA as template using an oligo(dT)18 primer. From each reaction 5 µl were used for the follow-up gene-specific amplification of the vNAR domain using the primers bamboo/nat_up and bamboo/nat_lo (**3.4**). Polymerase chain reaction (PCR) conditions were as follows: 94 °C for 2 min, 30 cycles of 30 s at 94 °C, 30 s at 55 °C and 40 s at 72 °C, followed by 72 °C for 7 min. PCR products were pooled, purified *via* the Wizard® SV Gel and PCR Clean-Up System (Promega) and the vNAR repertoire was analyzed using next-generation sequencing on 454 GS Junior System platforms (Roche) according to the manufacturer's instructions. In general, in a first PCR multiplex identifier sequences (MID sequences) were fused to the PCR product for the identification and assignment of vNAR sequences. In a subsequent emulsion PCR (emPCR Kit *(Lib-A)*) single DNA fragments were immobilized on beads *via* specific oligonucleotides and amplified in "micro-reactor" emulsion droplets achieving clonality. After loading of the beads onto a pico-titer plate device the sequencing reaction was performed. Data was analyzed using the GS Junior System software package v2.7, Jemboss (v1.5), jalview(v2) and muscle(v3.8.31). Deep sequencing was performed in cooperation with Andreas Christmann and Janine Becker (Technische Universität Darmstadt).

To gather information about the established vNAR library (**4.2.1**), the amino acid distribution of CDR3 was assessed *via* deep-sequencing. Evaluation was conducted akin to the analysis of the natural repertoire of the bamboo shark. However, in contrast to this, plasmid DNA of the vNAR library was isolated (**4.2.2**) and amplified using primer pair GR_up/GR_lo. PCR conditions were as follows: 94 °C for 2 min, 30 cycles of 30 s at 94 °C, 30 s at 55 °C and 40 s at 72 °C, followed by 72 °C for 7 min. PCR products were purified *via* the Wizard® SV Gel and PCR Clean-Up System (Promega) and the vNAR repertoire was analyzed using next-generation sequencing on 454 GS Junior System platforms (Roche) according to the manufacturer's instructions.

4.2 Yeast surface display library construction and establishment of randomized sublibraries for affinity maturation

4.2.1 CDR3-randomized library construction

Blood samples of three individuals of the bamboo shark (*C. plagiosum*) were harvested from the caudal vein and subsequently transferred into Tri® Reagent BD (Sigma Aldrich). Total RNA from three individuals were used as template for cDNA synthesis using the Omniscript® Reverse Transcription Kit (Qiagen) and OdT-Oligonucleotides. From each individual three reactions were processed in parallel, each containing approximately 2 µg of RNA. From each reaction 5 µl were used for the follow-up gene-specific amplification as mentioned above. PCR products were pooled and purified using the Wizard® SV Gel and PCR Clean-Up System (Promega) according to the manufacturer's protocol. The PCR amplified vNAR products were used as starting material for the semi-synthetic library construction.

The initial library was established in three consecutive PCR-steps (**Fig. 14**). For all reactions the conditions were: 94 °C for 2 min, 35 cycles of 30 s at 94 °C, 30 s at 55 °C and 40 s at 72 °C, followed by 72 °C for 7 min. Primer sequences are listed in section **3.4**. In the first reaction PCR amplified vNAR product from the natural repertoire was used as template with the primer combinations FR1/CDR1/Tyr_up and FR3_lo. The forward primer replaced Cys by Tyr and incorporated a marginal diversity within CDR1. The PCR product was purified *via* Wizard® SV Gel and PCR Clean-up System (Promega). Subsequently, the second PCR was performed to fully randomize CDR3 using the primer-combination FR1_up and CDR3/rand/FR4_lo. After purification by Wizard® SV Gel and PCR Clean-up System, the PCR product was used as a template for the third reaction using primers with overlaps up- and downstream of the *Nhe* I and *Bam* HI restriction sites of the pCT-plasmid, respectively (GR_up and GR_lo) and purified by Wizard® SV Gel and PCR Clean-up System.

The pCT vector (**3.2.1**) was digested with *Nhe* I and *Bam* HI. The reaction was performed in a volume of 100 µl. For this, approximately 20 µg plasmid were used as well as 100 U restriction enzyme, respectively, and incubated overnight at 37 °C. The reaction mixture was purified *via* Wizard® SV Gel and PCR Clean-up System.

Library was established *via* an improved yeast transformation method according to Benatuil *et al*. in a homologous recombination-based process referred to as plasmid gap repair.[181] For electroporation 1-2 µg of the digested plasmid and 6-8 µg of insert were used. Settings on the Gene Pulser® were as followed: 2.5 kV and 25 µF. Time constants ranged from 3.0 to 4.5 ms. Library size was calculated by dilution plating after 2 days. Yeast cells (EBY 100) were transferred into SD-CAA medium. Stocks were stored at -80 °C in yeast freezing solution. For yeast surface display cells were grown overnight at 30 °C in SD-CAA medium, transferred into SG-CAA medium and incubated for 1-2 days at 20 °C.

4.2.2 Establishment of randomized sublibraries for affinity maturation

For affinity maturation of the particular enriched antigen binding population, plasmid-DNA was isolated from yeast cells after sorting the cells *via* FACS and incubation at 30 °C for two days. To this end, 2 ml of yeast cell suspension were centrifuged (16100 g, room temperature, 1 min) and resuspended in 200 µl yeast lysis buffer. Chloroform and phenol were added in a volume of 100 µl, respectively, and cells were incubated at room temperature and 1000 rpm for approximately 10 min. After centrifugation (16100 g, 4 °C, 15 min) the supernatant was used as template for library establishment in a volume of 3 µl per PCR.

Sublibraries with totally randomized CDR1 were constructed in a 3-step PCR with the following conditions: 94 °C for 2 min, 35 cycles of 30 s at 94 °C, 30 s at 55 °C and 40 s at 72 °C, followed by 72 °C for 7 min. Akin to randomization of CDR3, in the first PCR reaction, the primer pair CDR1rand_up/GR_lo was used to randomize CDR1 followed by two consecutive PCR reactions with primer pairs FR1_up/GR_lo and GR_up/GR_lo, respectively. After each reaction step, the PCR product was purified *via* Wizard® SV Gel and PCR Clean-up System. Gap repair cloning and transformations were executed as described above (**4.2.1**).

4.2.3 Construction of a HV2-diversified library for the isolation of bi-specific IgNAR V domains

For the construction of a HV2-randomized yeast surface display library based on EpCAM-binding, affinity optimized single clone α-EpCAM-vNAR 5005, total DNA was isolated and used as starting material. For this, 2 ml of yeast cell suspension were centrifuged and plasmid DNA was isolated as aforementioned (**4.2.3**).

This particular library was established in a consecutive two-step splicing by overlap extension PCR. For the first PCR step, two reactions were carried out in parallel, each containing 3 µl isolated plasmid DNA as template. In one reaction, primer pair HV2_SOE_rand_up/pCT_Seq_lo was used. In the other reaction, primer pair pCT_Seq_up/HV2_SOE_lo was used. PCR conditions were as followed: 94 °C for 2 min, 35 cycles of 30 s at 94 °C, 30 s at 55 °C and 40 s at 72 °C, followed by 72 °C for 7 min. The respective PCR products were purified using Wizard® SV Gel and PCR Clean-up System. For the subsequent PCR, 1 µl of PCR product was used as template, respectively. After 6 cycles, primer pair pCT_Seq_up/pCT_Seq_lo was added. The resulting PCR product was purified *via* Wizard® SV Gel and PCR Clean-up System. Gap repair cloning and transformations were executed as described above.

4.3 Binding assays on the yeast surface and library screening for the isolation of target-specific vNAR molecules

4.3.1 Library analysis using yeast surface display

Flow cytometry was used to analyze presentation on the yeast surface and for single clone analysis. About 10^7 cells were labeled consecutively with anti-cMyc antibody (monoclonal, mouse, made in-house) or with anti-HA-tag antibody (polyclonal, rabbit, eBioscience, diluted 1:10), anti-mouse IgG biotin conjugate (goat, Sigma Aldrich, diluted 1:10 in PBS) or anti-rabbit-biotin (goat, Sigma Aldrich) and streptavidin-allophycocyanin conjugate (eBioscience, diluted 1:10) for at least 10 min on ice.

4.3.2 Single clone analysis using yeast surface display

For single clone analysis, vNAR-presenting cells were either incubated with biotinylated or His-tagged antigen for 15 min on ice and subsequently stained with Steptavidin-APC (diluted 1:10) or Penta-His Alexa Fluor 488 conjugate (Qiagen, diluted 1:30) for 10 min. When biotinylated antigen was used, analysis of surface presentation was conducted using anti-cMyc antibody and anti-mouse FITC conjugate (goat, Sigma Aldrich, diluted 1:10).

For biotinylation of target proteins, Sulfo-NHS-LC-Biotin (Thermo Fisher Scientific) was used as reagent. For this, a 20-fold molar excess of biotin reagent was used. All reactions were carried out in PBS for 2 hours on ice. To remove non-reacted biotin reagent, biotinylated protein was purified using protein desalting sping columns (Thermo Fisher Scientific) according to the manufacturer's instructions for buffer exchange.

Plasmid DNA from FACS-positive clones was isolated as aforementioned (**4.2.2**) and PCR product, established in a PCR using primer pair pCT-seq_up/pCT_seq_lo, was sent out for sequencing with either pCT-seq_up or pCT-seq_lo.

4.3.3 Affinity titration on the yeast surface for the calculation of the equilibrium dissociation constant

Affinities of isolated vNAR variants on yeast cells were determined as described [156,182]. Flow cytometric analysis was performed to determine cMyc normalized antigen binding. Hence, exclusively cells displaying the vNAR domain were included for K_D calculation. Therefore, yeast cells (approximately 10^6 cells for each antigen-concentration) were labeled with anti-cMyc antibody for at least 15 min. Afterwards, cells were washed using PBS and resuspended in the respective concentration of antigen as well as anti-mouse IgG biotin conjugate was added. When biotinylated antigen was used for affinity calculation, anti-mouse FITC conjugate was employed. Cells were

incubated for approximately 1 hour with the respective antigen-concentration. After washing, detection reagents were added for about 10 min, followed by an additional washing step (Penta-His Alexa Fluor 488 conjugate, Steptavidin-APC or Anti-Hu IgG (Fcγ-specific) PE for detection of antigen binding as well as anti-mouse FITC or Steptavidin-APC for the detection of surface presentation). In general, at least eight different antigen concentrations were used. For the final calculation, cMyc-normalized RFU were plotted against antigen concentration. EC_{50} was determined using equation $y = y_{min}+(y_{max}*x/(x+K_D))$.

4.3.4 Library screening using yeast surface display

Library screening for the isolation of antigen-specific vNAR molecules was performed on a fluorescence activated cell sorter (FACS) MoFlo® Cytometer with Cyclone sorting unit and analyzed *via* Summit® v4.3. For two-dimensional screening, cells were first labeled for detection of surface presentation *via* cMyc labeling. After washing with ice-cold PBS, cells were resuspended in PBS containing the desired concentration of antigen and incubated on ice for at least 30 min. Subsequently, antigen binding was detected using streptavidin-allophycocyanin conjugate, Anti-Hu IgG (Fcγ-specific) PE or Penta-His Alexa Fluor 488 conjugate (diluted 1:10, 1:30 or 1:30, respectively, 10 min on ice). For the first rounds of sorting of the initial library, approximately 2 x 10^8 cells were analyzed and sorted. Consecutive rounds were performed with at least a 10-fold excess of cells that were collected in the previous round to ensure coverage of the enriched population. Sublibrary screening was executed with at least 10-fold the number of cells of the theoretical library diversity**.**

4.4 Soluble expression of vNAR constructs in *E. coli* and protein purification

4.4.1 Cloning of vNAR fragments

Selected vNARs were expressed in the pMX vector [183] that introduces maltose binding protein (MalE) and a hexahistine tag for fusion protein expression and purification (**Fig. 26**).

pMX vector, containing for *E.*coli codon-optimized α-EpCAM-vNAR DNA-sequences for variant 5005 was purchased at GeneArt (Life Technologies). Codon-optimization was performed according the manufacturer's software. For expression of α-EpCAM-vNAR variants H3 and H5, total DNA was isolated from the respective single clones (**4.2.2**). IgNAR V domain fragments were amplified in a PCR reaction using primer combination pMX_vNAR_up/pMX_vNAR_lo and 3 µl isolated total DNA as template. PCR conditions were as followed: 94 °C for 2 min, 35 cycles of 30 s at 94 °C, 30 s at 55 °C and 40 s at 72 °C, followed by 72 °C for 7 min. PCR products were purified using Wizard® SV Gel and PCR Clean-up System.

Afterwards, PCR-products as well as pMX-vector (3.2.2) were digested using *Acc* 65I and *Xba* I. PCR products as well as pMX plasmid were digested in a volume of 50 µl, respectively. For this, 40 µl PCR product or pMX plasmid was used as well as 50 U restriction enzymes, respectively. Digestion was performed over night at 37 °C. Digested PCR products were purified *via* Wizard® SV Gel and PCR Clean-up System. Digested pMX plasmid was gel-purified (1 % w/v Agarose) also using the Wizard® SV Gel and PCR Clean-up System according to the manufacturer's protocol.

Subsequently digested vNAR DNA fragments were ligated into the pMX vector. For this purpose, 10 µl plasmid (corresponding to approximately 100 ng gel-purified plasmid) and 10 µl digested vNAR DNA (about 400 ng) were utilized. Reaction was performed in a volume of 30 µl with 400 U T4 DNA ligase added, for 2 hours at room temperature. After this, reaction mixture was purified using the Wizard® SV Gel and PCR Clean-up System.

Electro-competent *E. coli* TOP 10 cells (100 µl) were transformed with 15 µl purified ligation mixture. Settings on the Gene Pulser® were as followed: 2.5 kV and 25 µF. Time constants varied between 4.5 and 5.0 ms. For the generation of electro-competent TOP 10 cells, *E. coli* cells were grown to an OD_{600} of approximately 0.7 in a volume of 50 ml and washed thrice with ice-cold deionized water (50 ml, 30 ml, 20 ml). Cells were resuspended in the residual supernatant and used for transformation. Plasmid DNA of transformed TOP 10 cells was isolated using Wizard® Plus SV Minipreps DNA Purification System (Promega) according to the manufacturer's instructions and sent out for sequencing.

4.4.2 Soluble expression and purification

Electro-competent *E.coli* BMH 71-18 cells were generated akin to the protocol for electro-competent *E.coli* TOP 10 cells (**4.4.1**) and transformed with plasmids pMX_αEpCAM-vNAR_H3, pMX_αEpCAM-vNAR_H5 and pMX_αEpCAM-vNAR_5005, respectively. Cells containing the respective plasmid were grown to an OD_{600} between 0.7 and 1 in dYT medium (1 L) containing 25 mg/L chloramphenicol and protein production was induced with 1 mM IPTG. Cells were grown overnight (approx. 16 h) at 29 °C and harvested by centrifugation.

For the isolation of the periplasmic fraction, cells were resuspended in approximately 40 ml ice-cold periplasm buffer. 40 µl of a lysozyme stock (100 mg/ml in periplasm buffer) were added and incubated on a tumbling shaker for 30 min on ice. After centrifugation (16000 g, 15 min, 4 °C) the supernatant was harvested.

Respective recombinant vNAR proteins were purified by metal chelate affinity chromatography (HisTrap, GE Healthcare), equilibrated in PBS, pH 7.4. Protein was eluted using a linear imidazole gradient (0 % - 100 % of IMAC elution buffer over 30 min) and dialyzed overnight in 5 L PBS at 8 °C (ZelluTrans, Carl Roth GmbH, MWCO 3500).

For tabacco etch virus protease (TEV) cleavage, 10-fold TEV buffer (1.5 M NaCl, 500 mM Tris-HCl, pH 7.5) and recombinant TEV protease (ratio 1:20) were added and incubated overnight at 4 °C.

Final purification was performed using gel filtration on either a Superdex 200 pg 16/60 or on a Superdex 75 pg 16/60 (GE Healthcare), equilibrated with PBS, pH 7.4.

Proteins were analyzed *via* SDS-PAGE (15 % v/v SDS gels). For this, 5x PAGE loading buffer was added to protein samples. After electrophoresis, gel was stained using coomassie staining solution.

4.5 Soluble expression of formatted vNAR-Fc fusion constructs in mammalian cells and protein purification

4.5.1 Formatting of vNAR fragments

EpCAM-binding, affinity-optimized vNAR 5005 as well as its bi-specific progeny, EpCAM- and CD3ε-binding vNAR B1mut were expressed in mammalian cells. Therefore, both IgNAR V domains were cloned into pExpress vector that introduces the human Fc-part (**3.2.4**).

For this, total DNA was isolated from single clones, respectively (**4.2.2**). IgNAR V domain DNA fragment for mono-specific variant 5005 was amplified in a consecutive 2-step PCR containing primer pairs pEXPR_vNAR_up/pEXPR_vNAR_loI and pEXPR_vNAR_up/pEXPR_vNAR_loII, respectively. PCR conditions were as followed: 94 °C for 2 min, 35 cycles of 30 s at 94 °C, 30 s at 55 °C and 40 s at 72 °C, followed by at 72 °C for 7 min. First PCR was executed using 3 µl isolated total DNA. PCR product was purified PCR using Wizard® SV Gel and PCR Clean-up System. 1 µl PCR product was used as template for the second reaction.

Akin to this, vNAR DNA fragment for bi-specific variant B1mut was amplified under same conditions. However, primer pairs pEXPR_vNAR_up/pEXPR_B1mut_loI and pEXPR_vNAR_up/pEXPR_B1mut_loII were utilized.

Afterwards, PCR-product as well as pExpress-Fc plasmid (**3.2.4**) were digested using *Apa* I and *Nhe* I. PCR products as well as pExpress-Fc plasmid were digested in a volume of 50 µl, respectively. For this, 40 µl PCR product or pMX plasmid were at first incubated with 50 U *Apa* I at 25 °C for 3 hours. Then 50 U *Nhe* I were added and reaction mixture was incubated overnight at 37 °C. Digested products were purified *via* Wizard® SV Gel and PCR Clean-up System.

Subsequently digested vNAR DNA fragments were ligated into the pExpress-Fc vector. For this purpose, 5 µl plasmid (corresponding to approximately 300 ng plasmid) and 10 µl digested vNAR DNA (about 500 ng) were used. Reaction was performed in a volume of 30 µl with 400 U T4 DNA ligase added, for 2 hours at room temperature. After this, reaction mixture was purified using the Wizard® SV Gel and PCR Clean-up System.

Electro-competent *E. coli* TOP 10 cells (100 µl) were transformed with 15 µl purified ligation mixture. Settings on the Gene Pulser® were as followed: 2.5 kV and 25 µF. Time constants varied between 4.5 and 5.0 ms. Plasmid DNA of transformed *E. coli* TOP 10 cells was isolated using Wizard® Plus SV Minipreps DNA Purification System and sent out for sequencing. For transfection of mammalian cells, large amounts of plasmid DNA were generated and isolated using the Wizard® Plus SV Midipreps DNA Purification System (Promega) after the manufacturer's instructions.

4.5.2 Soluble expression of vNAR-Fc fusion proteins using the Expi293™ Expression System Kit

Expi293™ cells were cultured at 37 °C, 5 % CO_2 and 100 rpm in Expi293™ expression medium. Transfection was performed in a 30 ml scale. For this, 2.9 x 10^6 viable cells/ml were seeded in approximately 25 ml of Expi293™ expression medium. 30 µg plasmid DNA (pExpress_αEpCAM-vNAR_5005) were added to Opti-MEM® I reduced serum medium to a total volume of 1.5 ml. 81 µl of ExpiFectamine™ 293 were diluted in Opti-MEM® I reduced serum medium to a volume of 1.5 ml. After 5 min incubation, Opti-MEM® I reduced serum medium containing the DNA was added to the ExpiFectamine™ 293/Opti-MEM® I reduced serum medium mixture. After 20 min incubation at room temperature, transfection mixture was added to the cells. After one day incubation (37 °C, 5 % CO_2 and 100 rpm), 150 µl ExpiFectamine™ 293 transfection enhancer 1 and 1.5 ml ExpiFectamine™ 293 transfection enhancer 2 were added. Cells were incubated for 4 more days and supernatant was harvested by centrifugation (4500 g, 10 min, 4 °C). Soluble expression with Expi293™ cells was carried out at Merck KGaA (Darmstadt).

4.5.3 Soluble expression of vNAR-Fc fusion in HEK293 cells and purification

HEK293 cells were cultured at 37 °C, 5 % CO_2 and 100 rpm in Freestyle™ 293 expression medium. One day post transfection approximately 1 x 10^6 viable cells/ml were seeded in 25 ml Freestyle™ 293 expression medium. On the day of transfection, cells were typically grown to 1.7 – 1.9 x 10^6 viable cells/ml. Transfections were performed in a 25 ml scale. Therefore, 25 µg plasmid DNA were diluted in culture medium to a total volume of 1.25 ml and vortexed shortly. 75 µg PEI were diluted in culture medium to a volume of 1.25 ml and treated equally. Afterwards, PEI-culture medium mixture was added to the DNA-culture medium mixture, vortexted and incubated at room temperature for 15 min. Transfection mixture was added cautiously in droplets to the cells. After about 24 hours, 700 µl of a 20 % (w/v) trypton solution (in culture media) was added and cells were incubated for 4 more days at 37 °C, 5 % CO_2 and 100 rpm. Cells were incubated for four more days and supernatant was harvested by centrifugation (4500 g, 10 min, 4 °C).

Respective recombinant vNAR-Fc fusion proteins were purified by Protein A affinity chromatography using a HiTrap Protein A HP 1 ml column (GE Healthcare), equilibrated in Protein A chromatography running buffer, pH 7. For this, the supernatant was diluted with Protein A chromatography running buffer in a 2:1 ratio and injected onto the column. vNAR-Fc fusion protein was eluted using Protein A chromatography elution buffer, pH 3. To immediately neutralize acidic pH and thus, to minimize potential protein denaturation, 1 ml fractions were collected in tubes containing 200 µl of Protein A chromatography collection buffer, pH 9.0. Collected fractions were pooled and dialyzed (ZelluTrans, Carl Roth GmbH, MWCO 3500) against 5 L PBS, pH 7.4 for at least 3 hours at 8 °C and protein quality was assessed *via* SDS-PAGE (**4.2.2**).

4.6 Characterization of vNAR proteins

4.6.1 Determination of binding kinetics using the Octet RED96 System

Biolayer interferometry (BLI) was used to determine the equilibrium dissociation constant (K_D). All assays were performed on the Octet RED96 system (FortéBio). Affinities of target binding vNAR molecules were calculated with solitary vNAR domains, produced in *E. coli* as well as with vNAR-Fc fusion proteins, produced in human cell lines.

For affinity calculation of solitary vNAR domains, measurements were performed with Streptavidin dip and read biosensors in kinetics buffer (PBS, pH 7.4, 0.1 % (w/v) BSA, 0.02 % Tween-20). Streptavidin dip and read biosensors were rehydrated for at least 30 min in PBS and put into the Octet Red96 device. Sensors were loaded with biotinylated vNARs at approx. 10 µg/ml. A target-unspecific vNAR was used as negative control and tested at the highest antigen concentration. For all measurements, kinetic data sets were fitted using 1:1 Langmuir binding *via* the manufacturer's data analysis software with Savitzky-Golay filtering.

For affinity assessment of vNAR-Fc fusions, measurements were performed with anti-human IgG Fc capture dip and read biosensors in kinetics buffer. Biosensors were rehydrated for at least 30 min in PBS and put into the Octet Red device. Sensors were loaded with Fc fusions at approx. 10 µg/ml. As negative controls, unspecific binding of target proteins to the biosensor tips was tested as well as unspecific binding of the immobilized vNAR-Fc fusion to an unrelated target protein.

4.6.2 Thermal shift assays

Thermal shift assays were conducted to determine the Tm. Measurements were performed in quadruplicates on a BioRad 96CFX RT-PCR detection system with 0.5 °C/30 s to 99 °C. Tm values were obtained from melting curves using the corresponding BioRad analysis software. All reactions were performed in PBS, pH 7.4 in the presence of SYPRO Orange (diluted 1:1000, Sigma-Aldrich) and 0.1 mg/ml – 0.5 mg/ml protein.

4.6.3 Cellular immunofluorescence assays

Trypsinized cell lines MCF-7, EBC-1, T47D as well as CHO(s) were generously provided by Merck (EMD) Serono. For each antibody or vNAR-Fc concentration tested, approximately 5 x 10^5 cells of the respective cell line were centrifuged and used for each antibody and vNAR staining. Assays were performed in a volume of 100 µl in PBS, pH 7.4 containing 1 % (w/v) BSA. Cells were first centrifuged (250 g, 4 °C, 10 min), and resuspended in different concentrations of α-EpCAM-vNAR 5005 Fc fusion protein. Cells were stained for approximately 1 hour on ice. EpCAM-specific antibody clone HEA-125 conjugated to PE (MACS, Miltenyi Biotec) was used as positive control in a dilution of 1:11. After washing the cells with ice-cold PBS, pH 7.4 containing 1 % (w/v) BSA, cells were incubated with anti-human IgG conjugated to PE (Fcγ-specific, eBioscience, diluted 1:30), except for EpCAM-specific antibody HEA-125. Cells were analyzed *via* FACS.

5. RESULTS

5.1 Analysis of the natural vNAR repertoire of the bamboo shark (*C. plagiosum*)

In a previous work tissues as well as blood samples were harvested from anesthetized bamboo sharks.[184] After total RNA isolation as well as reverse transcription, tissue-specific expression of IgNAR was analyzed *via* vNAR-specific amplification. It revealed that besides expression in blood, IgNAR is also found on the mRNA-level in the rectal gland, the skin, the liver, the stomach, the spleen, the kidney as well as in the gonads and in the epigonal organ.[184]

To further gather information about the IgNAR V domain repertoire from non-immunized *C. plagiosum*, the diversity of the natural repertoire as well as the length distribution of CDR3 was examined (**4.1**). All respective deep sequencing subprojects within this work were carried out in cooperation with Andreas Christmann and Janine Becker (Technische Universität Darmstadt). For this, total RNA was isolated from whole blood of one specimen of the bamboo shark. Reverse transcription was executed using an Oligo(dT)18 primer. After gene-specific amplification using the primer combination bamboo/nat_up and bamboo/nat_lo, PCR products were pooled and the vNAR repertoire was analyzed using next-generation sequencing on the 454 GS Junior System platform (Roche) according to the manufacturer's instructions. Consistent with literature,[109] deep-sequencing of the vNAR repertoire from blood circulating lymphocytes of one individual of *C. plagiosum* revealed that IgNAR V region diversity in this species is mainly CDR3 based (**Fig. 13A**). CDR1 along with HV2 and HV4 only display minor sequence variations. The vast majority of approximately 92% of analyzed vNAR sequences had one cysteine residue within CDR3 and within CDR1, respectively, potentially forming an inter-loop disulfide linkage. Hence, the vNAR repertoire of the bamboo shark is considered to be mainly of type II, according to the categorization of vNAR domains based on the presence and absence of non-canonical disulfide bonds (**2.5.2**). As shown in **Figure 13B**, most of analyzed clones had a CDR3 with a length of 12 residues. However, a considerable fraction of clones contained a significantly prolonged length of 16 or more residues. A mean length of about 13 residues (12.8 aa) is substantially elongated compared to murine CDR3 sequences of the VH domain with a mean length of about 9 residues.[185] Surprisingly, the human CDR-H3 repertoire displays a broader length distribution from 1 to 35 residues compared to the bamboo shark (5-22 residues).[186] Still, the mean length of about 12 residues[185] is comparable with the mean length of CDR3 of the vNAR domain of *C. plagiosum.*

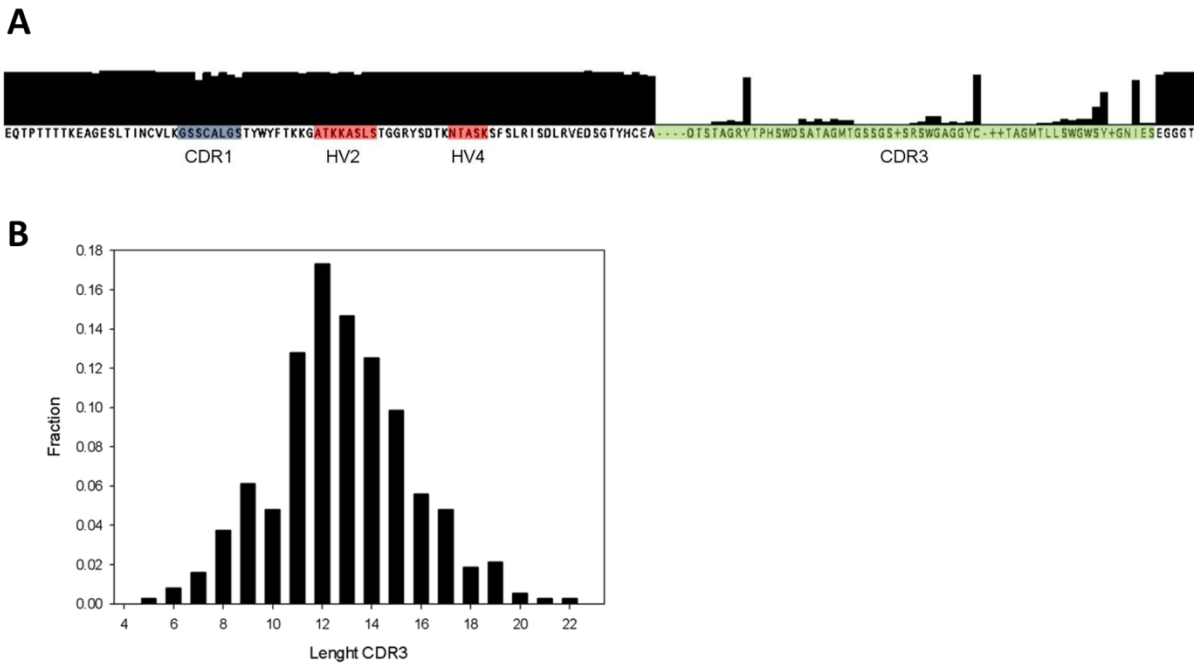


Figure 13. Analysis of the natural vNAR repertoire of the bamboo shark. (A) Consensus-sequence of the natural vNAR repertoire from non-immunized bamboo sharks. A set of approx. 1200 sequences was analyzed. The primer-mediated parts at the respective *N*-terminus and *C*-terminus are not shown. CDR1 is shaded blue, HV2 and HV4 are shaded red, and CDR3 is shaded green. Consensus of residues at each position is indicated by a black bar. (B) Length distribution of CDR3 of the vNAR repertoire from a non-immunized bamboo shark.

5.2 Construction of a semi-synthetic type IV bamboo shark vNAR library for yeast surface display

Parts of this work have been published.[187] Results gained from the analysis of the primary vNAR repertoire of the bamboo shark had several consequences for the construction of the IgNAR V domain library for yeast surface display. We set out to construct a semi-synthetic type IV library using the natural PCR-amplified framework repertoire of *C. plagiosum* as starting material. Library construction on *S. cerevisiae* allows for the facile generation of more than 10^8 clones. Taking this into consideration, we assumed that a total randomization of CDR3 using codon-based oligonucleotides almost certainly exceeds the natural diversity found in non-immunized animals. As a consequence, to extend natural vNAR diversity, we constructed the semi-synthetic vNAR library by polymerase chain reaction where in the framework of the natural vNAR repertoire CDR3 was totally randomized. This was achieved by the incorporation of trinucleotide mixtures encoding 19 amino acids except cysteine into the corresponding oligonucleotide. Framework regions and the hypervariable loops were used from non-immunized repertoires of a cohort of three bamboo sharks. In analogy to the natural repertoire (**Fig. 13B**) we designed the CDR3 to comprise 12 randomized residues.

For library construction total RNA was isolated from whole blood of three individuals and used as template for cDNA synthesis using the Omniscript® Reverse Transcription Kit (Qiagen) and OdT-

Oligonucleotides (**4.2.1**). From each individual three reactions were processed in parallel, each containing approximately 2 µg of RNA. From each reaction 5 µl were used for the follow-up gene-specific amplification. The pooled PCR amplified naïve vNAR products were used as starting material for the semi-synthetic library construction. The initial library was established in three consecutive PCR-steps (**Fig. 14B**). In the first reaction PCR amplified vNAR product from the natural repertoire was used as template (approximately 200 ng) with the primer combinations FR1/CDR1/Tyr_up and FR3_lo. The forward primer replaced Cys by Tyr and incorporated a marginal diversity within CDR1 to resemble the natural sequence variation found in this region (**Fig. 13A**). Subsequently, the second PCR was performed to fully randomize CDR3 using the primer-combination FR1_up and CDR3/rand/FR4_lo. Afterwards, this particular PCR product was used as a template for the third reaction using primers with overlaps up- and downstream of the *Nhe* I and *Bam* HI restriction sites of the pCT-plasmid[150], respectively (GR_up and GR_lo). Accordingly, the pCT vector was digested with *Nhe* I and *Bam* HI. Library was established *via* an improved yeast transformation method according to Benatuil *et al*. in a homologous recombination-based process referred to as plasmid gap repair.[181] For electroporation 1-2 µg of the digested plasmid and 6-8 µg of insert were used.

Two days after transformation, the library size was calculated by dilution plating and yielded in a library diversity of approximately 2×10^8 unique clones. Library-encoded IgNAR V domains were displayed on the yeast cell surface *via* fusion to cell wall protein Aga2p, which is anchored to the cell wall by association with Aga1p.[150] The vNAR molecules were flanked by a HA-tag and cMyc-tag at the *N*-terminus and *C*-terminus, respectively, for detection of surface-exposed vNAR (**Fig. 14A**). Surface presentation of vNAR variants (**4.3.1**) was characterized by indirect fluorescence labeling of the cMyc-epitope and HA-tag (**Fig. 14C** and **D**). Within three days post-induction, there was nearly total HA-tag expression of the library detectable and nearly 90% of cMyc-labeling, both indicating expression and high copy number display of vNAR domains on the surface of *S. cerevisiae*.

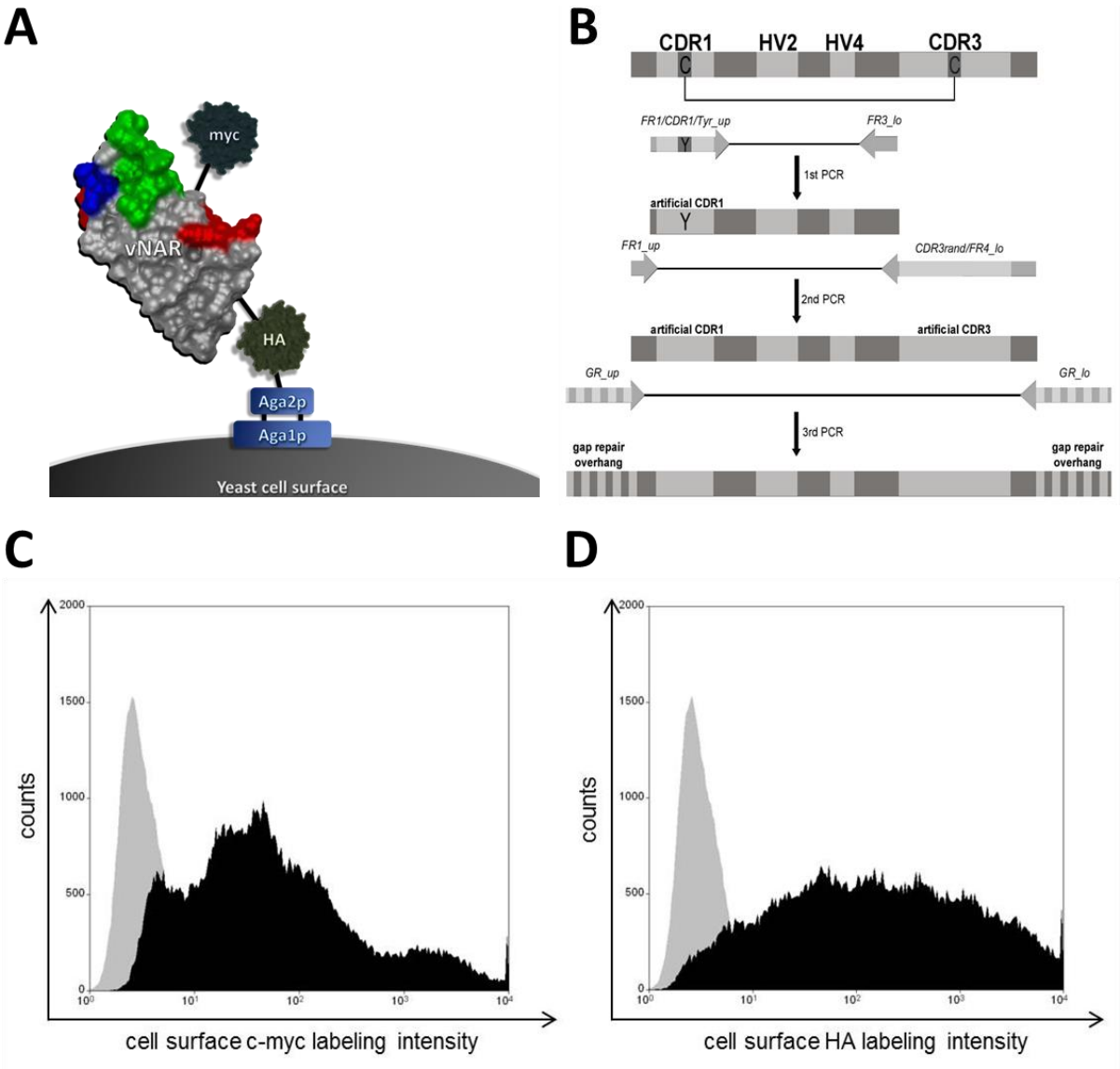


Figure 14. Library construction and analysis. (A) Illustration of vNAR yeast surface display.[149] vNARs are presented on the surface of *S. cerevisiae* as Aga2p fusions with an *N*-terminal HA-epitope and a *C*-terminal cMyc-tag, respectively, for the detection of surface expression. (B) Schematic representation of PCR-based library design. PCR amplified vNAR fragments from blood circulating lymphocytes of a cohort of three bamboo sharks were used as template. In a first PCR the framework was amplified. The cysteine residue in CDR1 was replaced by tyrosine and a marginal diversity was introduced *via* the forward primer. In a subsequent reaction the total vNAR molecule was constructed and an artificial CDR3 was introduced. The 3rd reaction was carried out to incorporate homologous sites for gap repair cloning. Histogram of cMyc (C) and HA-tag (D) surface expression of the constructed vNAR library assessed by indirect immunofluorescence labeling and flow cytometry. Grey: negative control (cells after 3 days induction of gene expression) labeled with detection reagents only; black: immunofluorescence staining after 3 days induction of gene expression.

To further evaluate library quality, the distribution of amino acids and the codon usage in the artificial CDR3 were assessed *via* deep sequencing (**4.1**, **Fig. 15**). 8220 positions from 685 unique vNAR sequences with a CDR3 of 12 residues were analyzed on the 454 GS Junior System platform (Roche) according to the manufacturer's instructions. The horizontal bar shown in **Figure 15** depicts the distribution of 5.26 % for each amino acid residue in an ideally unbiased library (**Fig. 15**). The residue

overrepresented by more than 2 % was Lys with a frequency of 8.37 %. The residues underrepresented in this library by more than 2 % were Phe and Val with frequencies of 3.24 % and 3.08 %, respectively. Albeit not considered for library design, Cys residues and stop codons were present in CDR3 with frequencies of 0.49 % and 0.1 %. Those findings may originate from mistakes during deep sequencing procedures or to a marginal fraction of wild-type vNAR CDR3.

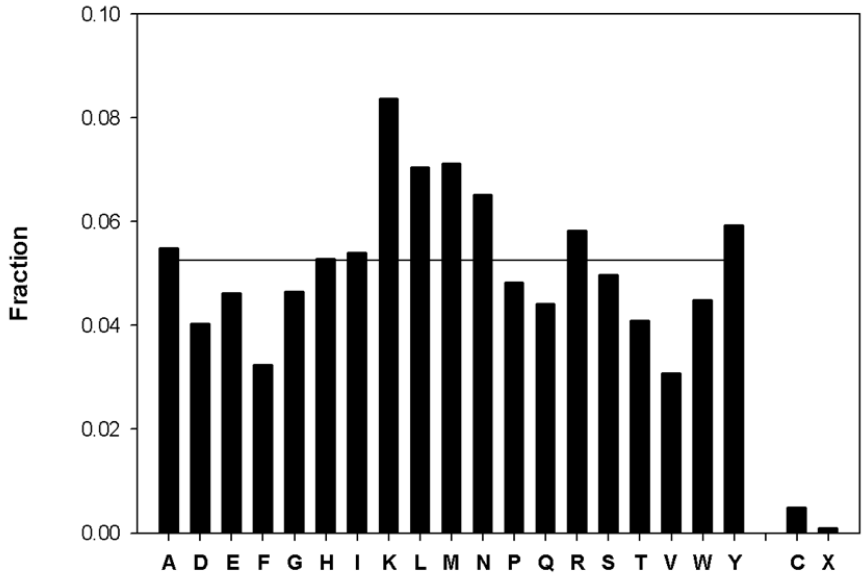


Figure 15. Distribution of amino acids in CDR3 of the unselected library. A total number of 685 unique sequences were analyzed (8220 positions). X, Stop-codon. The horizontal bar marks the expected representation for all 19 residues of an ideally unbiased library.

5.3 Selection of antigen binding semi-synthetic vNAR domains

The established vNAR library was subjected for library screening to select for mono-specific target binding molecules (**4.3.4**). Screening was conducted against EpCAM, EphA2 and Interleukin-8 (the latter in cooperation with Julius Grzeschik during his master-thesis). Additionally, antigen binding vNAR molecules were also selected to target human serine protease HTRA1, executed by Niklas Weber. Detailed data for the HTRA1 screening will be provided by Niklas Weber in his PhD thesis and can also be found elsewhere.[187]

5.3.1 Screening against epithelial cell adhesion molecule (EpCAM)

Recombinant human EpCAM (CD326) was purchased from AcroBiosystems. The protein consists of the EpEX domain (amino acids Gln 24 – Lys 265) fused to a poly-histidine tag at the *C*-terminus. The molecular weight is calculated to be 29 kDa. Due to glycosylation it migrates at 33-36 kDa.

All initial FACS sorting rounds were performed with 1 µM EpCAM. Target binding was detected by indirect immunofluorescence using either Alexa Fluor 488 labeled anti-His-tag antibody or using

biotinylated target protein followed by cell staining with streptavidin-allophycocyanin conjugate. All screening rounds were performed to simultaneously select for full-length vNAR presentation using myc-staining and for target binding. In the first round of sorting approximately 2 x 10^8 cells were screened (**Fig. 16, left)**. Antigen binding was monitored *via* Alexa Fluor 488 labeled anti-His-tag antibody. After each sorting round, collected cells were immediately subjected for a resort to enhance stringency. To avoid off-target binding against detection reagents the labeling strategy was alternated from Penta-His antibody to biotinylated EpCAM after the first round of screening. Subsequent library screening in sorting rounds two and three revealed a significant enrichment of antigen binding populations (**Fig. 16, middle and right**).

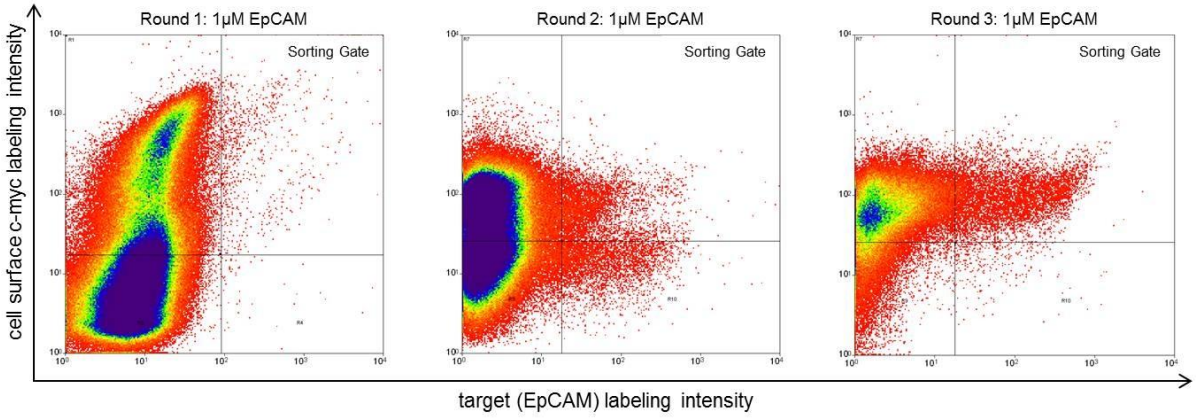


Figure 16. Library screening against EpCAM. Sorting gates and target concentrations are shown. Yeast cells were labeled for simultaneous detection of surface presentation and antigen binding. After each round a resort was performed. Cells in the sorting gate were isolated, grown and induced for the next round of selection. Percentage of cells in sorting gate: (R1) 0.17 %; (R2) 1.03 % and (R3) 8.04 %.

Single clones were analyzed for antigen binding (**4.3.2**) using 1 µM EpCAM an FACS-positive clones were sent out for sequencing. Ten unique EpCAM-binding vNAR clones were identified (**Fig. 17**).

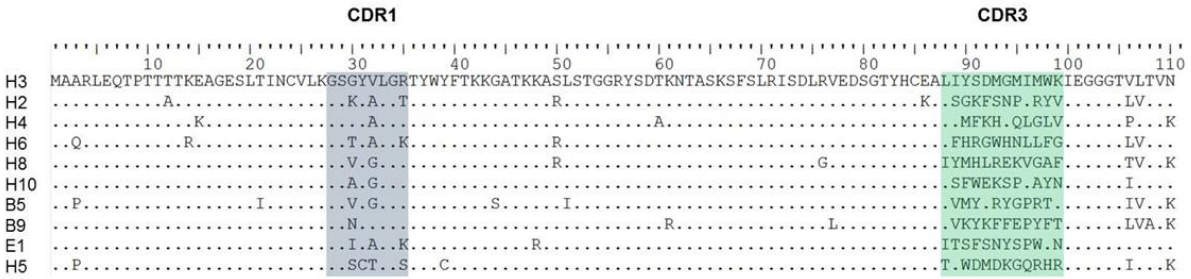


Figure 17. Sequences of EpCAM binding vNAR molecules selected from initial library screen. Dots indicate identical residues. CDR1 and CDR3 are shaded in blue and green, respectively.

Five clones were randomly chosen and more deeply characterized in terms of affinity *via* equilibrium dissociation constant determination on the yeast surface (**4.3.3**, **Tab. 2**). Affinity titration revealed

equilibrium binding constants for EpCAM binding clones in the triple-digit nanomolar to single-digit micromolar range, with an average affinity of 0.7 µM.

Table 2. Equilibrium dissociation constants (K_D) determined by yeast surface display.

vNAR clone	target	screen	K_D [nM]
H2	EpCAM	initial	741 ± 267
H3	EpCAM	initial	1301 ± 616
H5	EpCAM	initial	1030 ± 524
H8	EpCAM	initial	212 ± 161
H10	EpCAM	initial	301 ± 187
mmE1	EphA2	initial	368 + 144
mmE2	EphA2	initial	482 + 334
I1	IL-8	initial	> 3300

5.3.2 Screening against receptor tyrosine kinase EphA2

Recombinant human receptor tyrosine kinase EphA2 was purchased from R&D Systems. The protein consists of amino acids Gln25 - Asn534 and is fused to a *C*-terminal his-epitope. It has a predicted molecular mass of 56.9 kDa. Recombinant murine EpCAM was purchased from Sino Biological Inc. The extracellular domain (Met1 - Asn535) is fused to a His-eptitope at the *C*-terminus and has a predicted molecular weight of 58 kDa.

All screenings (**4.3.4**) were performed using 1 µM of EphA2. In all rounds we simultaneously selected for full-length vNAR presentation using myc-staining and for target binding. Akin to screening against EpCAM, after each round of screening, a resort was carried out. In the first round of screening, approximately 1.5×10^8 cells were sorted using indirect immunofluorescence *via* Alexa Fluor 488 labeled anti-His-tag antibody (**Fig. 18, upper row**). After sorting round 1 biotinylated target protein was used for consecutive screening rounds. In screening round 3, a target binding population was

detected. Consequently, single clones were analyzed for antigen binding. EphA2-binding clones were sent out for sequencing and four different EphA2-specific vNAR were identified (**Fig. 19**).

For the isolation of IgNAR V domains that bind both human and murine EphA2, which may be useful for animal studies, the population enriched for EphA2 binders (after sort 3) was subjected to two more rounds of selection *via* YSD using murine EphA2 as antigen (**Fig 18, lower row**). Surprisingly, in spite of alternated screening from his-tagged EphA2 and Alexa Fluor 488 labeled anti-His-tag antibody to biotinylated antigen, a distinct population of anti-His-tag antibody binding clones was detected in round 4 (**Appendix A**). To avoid the isolation of this particular population, the sorting gate was accordingly adjusted. The resort of this round showed that there was no significant enrichment of off-target binding clones (**Fig. 18, lower row, middle**). Along with this finding, in the fifth round biotinylated murine EphA2 was used for screening. After this round, single clones were analyzed for binding murine and human EphA2 (**4.3.2**). This resulted in two vNARs with the desired promiscuous specificity (**Fig. 19**). Affinities for both, human and murine EphA2 binding clones were characterized *via* affinity-titration on the yeast surface (**4.3.3**) using human EphA2 as antigen, resulting in an average affinity of around 0.4 µM (**Tab. 2**).

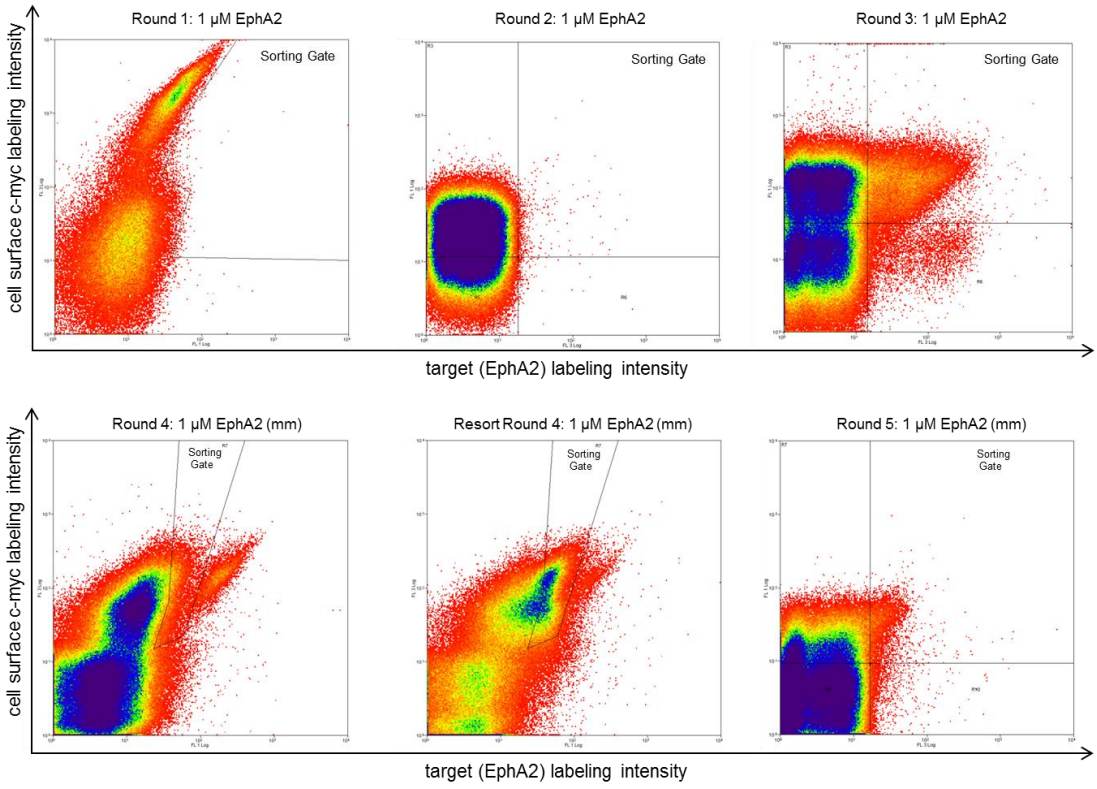


Figure 18. Library screening against EphA2. Sorting gates and target concentrations are shown. Yeast cells were labeled for simultaneous detection of surface presentation and antigen binding. After each round a resort was performed. Cells in the sorting gate were isolated, grown and induced for the next round of selection. Percentage of cells in sorting gate: (R1) 0.11 %; (R2) 0.10 % and (R3) 6.87 %. In round 4 and 5 murine EphA2 (mm) was used as antigen. Percentage of cells in sorting gate: (R4) 2.43 % and (R5) 1.01 %.

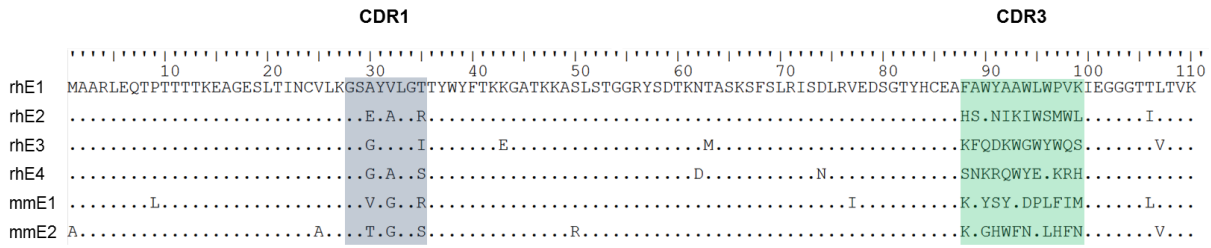


Figure 19. Sequences of EphA2 binding vNAR molecules selected from initial library screen. Dots indicate identical residues. CDR1 and CDR3 are shaded in blue and green, respectively. rh: vNAR molecules selected from screen against recombinant human EphA2. mm: clones selected after two additional rounds of screening against murine EphA2, promiscuous specificity for human and murine EphA2 was validated *via* yeast surface display.

5.3.3 Screening against human interleukin-8

The isolation (and affinity maturation) of Interleukin-8 binding vNARs was carried out in cooperation with Julius Grzeschik during his Masterthesis. Detailed data on the isolation and characterization can be found elsewhere.[188] Yet, main findings will be described within this subsection.

Interleukin-8 was provided by the working group of Prof. Schmitz, TU Darmstadt. The 8.3 kDa molecule was expressed in *E. coli* and purified by anion exchange chromatography, followed by size exclusion chromatography and HPLC on a RP-8 column. Target protein quality was assessed *via* SDS-PAGE and revealed no proteolytic degradation of IL-8 and no significant impurities. Since the provided IL-8 did not have any affinity-tag or epitope for the detection *via* indirect immunofluorescence, only biotinylated IL-8 was used for screening experiments. However, to avoid off-target binding, detection of target binding was alternated from streptavidin-APC (or streptavidin-PE) to avidin-FITC. Both detection reagents display high-affinity binding to biotin, though sequence identity is very low, minimizing the chance for the enrichment of binders against secondary detection reagents.

For the isolation of IL-8 binding clones, 2 µM was used for every screening round (**4.3.4**). In the first screening round approximately 4 x 10^8 cells were sorted. Labeling for antigen binding was carried out using avidin-FITC and alternated to streptavidin-APC for the second and third screening round. The first three screening rounds were performed *via* double-labeling. In the fourth round, double-labeling for simultaneous detection of full-length presentations and antigen binding was omitted since cMyc surface expression was about 75 %. As shown in **Figure 20** we were able to enrich for CXCL8 binding vNAR populations within four rounds of screening (**Fig. 20**). Single clones were analyzed for antigen binding (**4.3.2**) and FACS-positive clones were sent out for sequencing, resulting in two unique sequences (**Fig. 21**). Since CDR3 of single clone I2 mainly comprised hydrophobic amino acids, this clone was considered to be unspecific and not further scrutinized. Single clone I1 was characterized in terms of affinity on the yeast surface (**4.3.3**) which resulted in a moderate equilibrium binding constant in the single-digit micromolar range (**Tab. 2**).

To this end, binders isolated against all different targets (including HTRA1, not shown) displayed low to moderate affinities in the triple-digit nanomolar to single-digit micromolar range (for HTRA1 also in the double-digit micromolar range). To improve affinities substantially, a new methodology for affinity maturation was employed, as described in the following section.

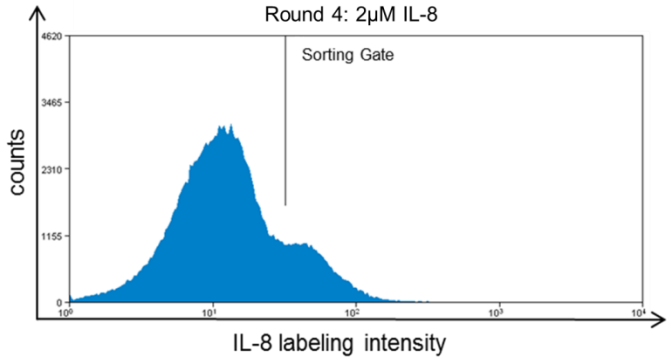


Figure 20. Fourth round of library screening against IL-8. Sorting gate and target concentration are shown. Yeast cells were labeled for antigen binding. Percentage of cells in sorting gate: 9 %.

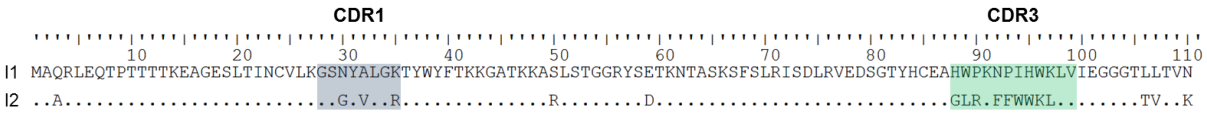


Figure 21. Sequences of IL-8 binding vNAR molecules selected from initial library screen. Dots indictate identical residues. CDR1 and CDR3 are shaded in blue and green, respectively.

5.4 Affinity maturation of target-enriched semi-synthetic vNAR domains

To enhance affinities of the selected antigen binding vNAR populations, a new methodology for affinity maturation was established that resembles the natural immune response in sharks. As mentioned in the introduction (**2.2.3, 2.5.3**), the primary vNAR repertoire of sharks is mainly restricted to CDR3. However, due to somatic hypermutation, after antigen contact, the vNAR domain further diversifies with mutations clustering to the CDRs and hypervariable loops. Furthermore a significant bias towards replacement substitutions is observed in CDR1.[109] To mimic this natural process, we decided to diversify five residues within CDR1 of vNAR molecules enriched after FACS-screening by constructing sublibraries. The design of these sublibraries for affinity maturation was based on the enriched target binding vNAR populations against EpCAM and HTRA1 (not shown within the scope of this thesis), respectively, which were obtained after three screening rounds. For Interleukin-8, only clone I1 was used as starting material for sublibrary design. As a consequence, for each target a particular sublibrary was constructed (**4.2.2**).

To this end, DNA was isolated from the enriched population or from the respective single clone and used as a template for the randomization of five residues in CDR1. Akin to randomization of CDR3 of

the initial library, diversifications were performed by incorporation of triplet codons into the corresponding oligonucleotide. For affinity maturation of the enriched antigen binding population or single clone plasmid-DNA was isolated from yeast cells after sorting the cells *via* FACS and incubation at 30 °C for two days. Sublibraries with totally randomized CDR1 were constructed in a consecutive 3-step PCR. Randomization of CDR1 was similar to the initial library establishment (**Fig. 14A**). In the first PCR reaction, the primer pair CDR1rand_up/GR_lo was used to randomize CDR1 followed by two subsequent PCR reactions with primer pairs FR1_up/GR_lo and GR_up/GR_lo, respectively. For affinity maturation of α-EpCAM-vNARs and α-IL-8-vNARs, libraries for yeast surface display were established according to initial library construction, as described in **chapter 5.2**. This yielded in a library diversity of approximately 2 x 10^8 unique clones for α-EpCAM-sublibrary and about 10^8 clones for α-IL-8-sublibrary, as determined by dilution plating. Since the theoretical diversity in a total randomized loop of five residues is approximately 2.5 x 10^6, library construction for both sublibraries was considered to almost ensure full coverage of all theoretical variants.

5.4.1 Sublibrary screening for the isolation of EpCAM binding vNARs with enhanced affinities

To obtain target binders with higher affinities, the EpCAM-sublibrary was subjected to FACS-screening using yeast surface display with significantly reduced target concentrations (**4.3.4**). Whereas the initial library was screened using 1 µM of EpCAM, in the first round of sublibrary screening only 50 nM of target was used (**Fig. 22**). In the second round of sorting, target concentration was once more step-wise reduced, aimed at enhancing selection stringency. Moreover, staining was alternated from Alexa Fluor 488 labeled anti-His-tag antibody staining to biotinylated antigen, to avoid enrichment of off-target binders. To isolate a multitude of affinity-enhanced binders and to avoid out-competition of binders, the target concentration was increased to 50 nM for α-EpCAM CDR1 affinity maturation in the last screening round. As shown in **Figure 18** a target binding population under enhanced selection stringency was enriched after round 2 and approximately 47 % target-positive cells were enriched in the last screening round (**Fig. 22**). Single clone analysis (**4.3.2**) and subsequent sequencing of target-positive variants revealed six different EpCAM-binders (**Fig. 23**). Interestingly, EpCAM-binding clones 1017, 5005 and R23 contained the same CDR3 as EpCAM-binding clone H3 from initial screenings. However, CDR1 was unique in all of those clones, respectively. In addition, clones 1013, 1014 and R31 showed novel and unique CDR3 loops, which have not been identified after sequencing of binders from the initial library.

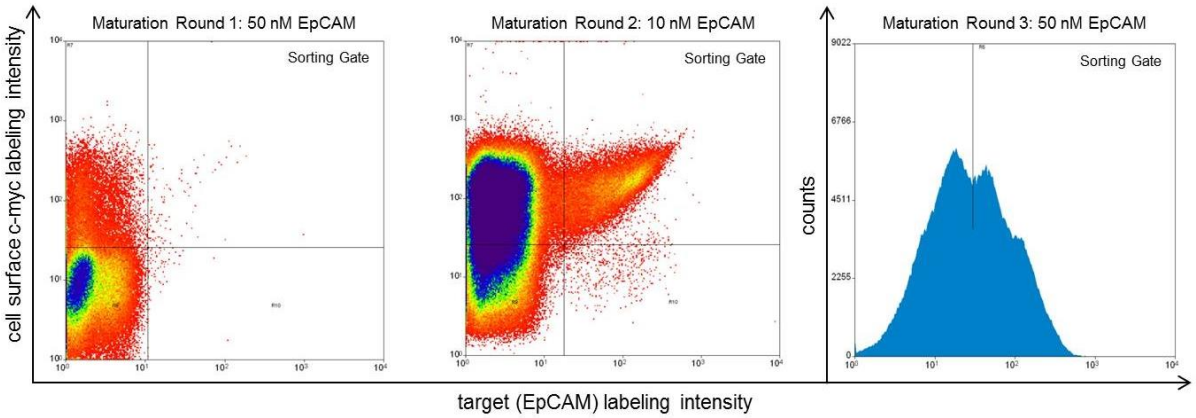


Figure 22. Screening of an α-EpCAM sublibrary after randomization of CDR1. Sorting gates and target concentrations are shown. Yeast cells were labeled for simultaneous detection of surface presentation and antigen binding (in Maturation Round 3, double labeling was omitted since vNAR cMyc surface expression was > 90%). After each round a resort was performed. Cells in the sorting gate were isolated, grown and induced for the next round of selection. Target-positive cells after affinity maturation in sorting gate: (MR1) 0.1 %; (MR2) 4.8 % and (MR3) 47.2 %.

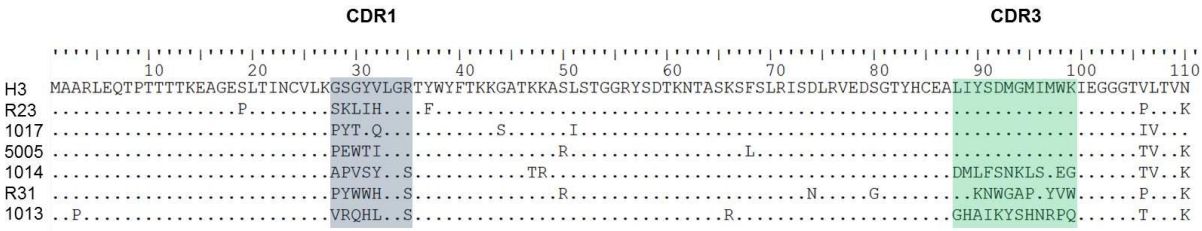


Figure 23. Sequences of selected α-EpCAM-vNARs after CDR1 randomization aligned to initially selected vNAR H3. Dots indicate identical residues. CDR1 and CDR3 are shaded in blue and green, respectively.

Five out of six clones against EpCAM after CDR1 affinity maturation were randomly selected for yeast cell surface affinity titration (**4.3.3**, **Tab. 3**). All EpCAM binders exhibited high-affinity binding in the low nanomolar range. The average affinity was 11.2 nM, which corresponds to an enhancement of affinity by the factor of 65 compared to the clones obtained prior to affinity maturation (**Tab. 2**). These findings clearly highlight the contribution of CDR1 to antigen binding.

5.4.2 Sublibrary screening for the isolation of CXCL8 binding vNARs with enhanced affinities

The established α-IL-8 sublibrary was screened, aimed at enriching CDR1-diversified variants with significantly enhanced affinities (**4.3.4**). For this, the target protein was step-wise reduced within successive screening rounds (**Fig. 24**). In this particular case, due to starting with one single-clone for sublibrary design and not with a whole population, there was no need for alternated screening, since enrichment of off-target binding clones against detection reagents could be excluded. Hence, in this particular affinity maturation procedure, antigen binding was exclusively detected using streptavidin-

APC. To maximize selection stringency the fourth round of sorting was performed in two parallel screening experiments with 25 nM and 10 nM, respectively (**Fig. 24, lower row**).

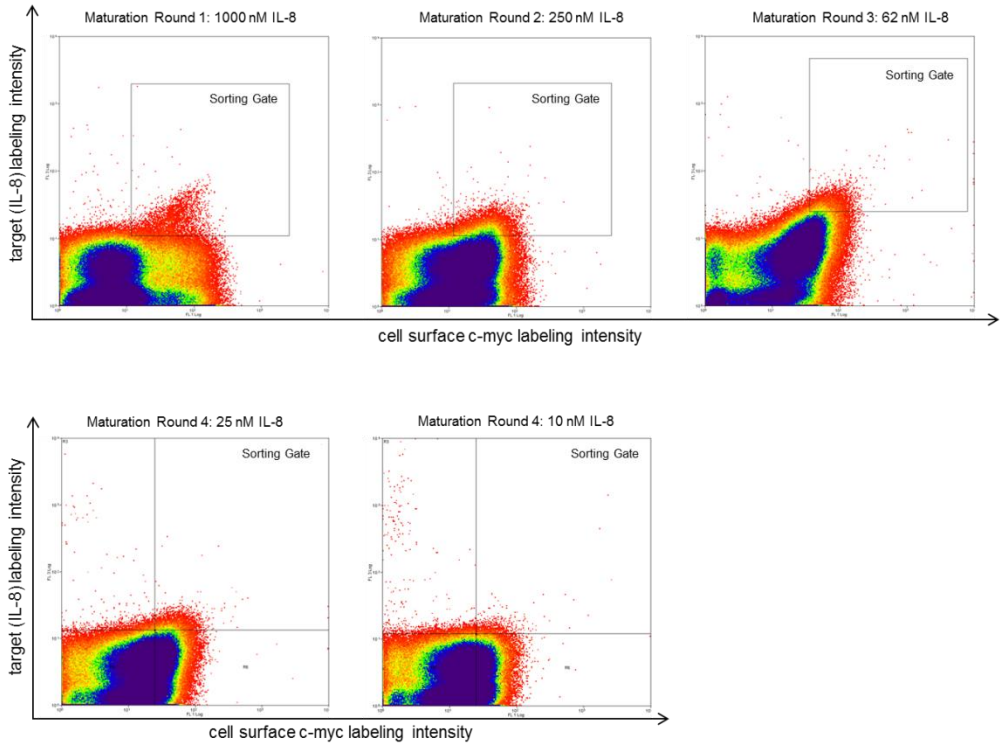


Figure 24. Screening of an α-IL-8 sublibrary after randomization of CDR1. Sorting gates and target concentrations are shown. Yeast cells were labeled for simultaneous detection of surface presentation and antigen binding. Cells in the sorting gate were isolated, grown and induced for the next round of selection. Target-positive cells after affinity maturation in sorting gate: (MR1) 0.8 %; (MR2) 2.5 %; (MR3) 1 %; (MR4, 25 nM) 0.93 % and (MR4, 10 nM) 0.24 %.

After the respective fourth round, single clones were analyzed for antigen binding (**4.3.2**) and FACS-positive clones were sent out for sequencing, which resulted in the identification of four vNARs, all comprising CDR3 of the parental molecule I1, but a unique CDR1 (**Fig. 25**). Affinities of all clones after affinity maturation were determined on the surface of the yeast *via* affinity titration (**4.3.3**) and exhibited significantly enhanced affinities compared to the parental molecules (**Tab. 3**). The best clone, referred to as I25.3, showed an affinity of about 430 nM, which corresponds to an enhancement of affinity by the factor of 7.6.

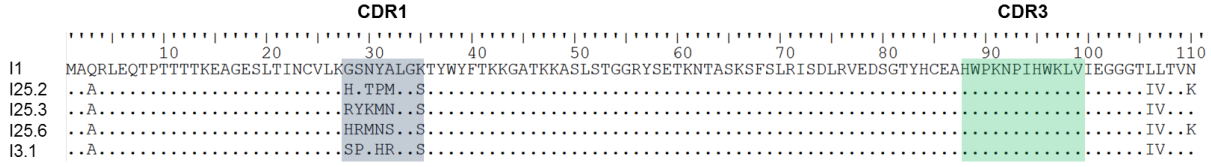


Figure 25. Sequences of selected α-IL-8-vNARs after CDR1 randomization aligned to initially selected vNAR I1. Dots indicate identical residues. CDR1 and CDR3 are shaded in blue and green, respectively.

Table 3. Equilibrium dissociation constants (K_D) determined by yeast surface display after affinity maturation.

vNAR clone	target	screen	parental CDR3	K_D [nM]
1013	EpCAM	CDR1mat	n.d.	7 ± 3
1014	EpCAM	CDR1mat	n.d.	5 ± 2
1017	EpCAM	CDR1mat	H3	6 ± 3
5005	EpCAM	CDR1mat	H3	14 ± 7
R23	EpCAM	CDR1mat	H3	24 ± 9
I25.2	IL-8	CDR1mat	I1	520 ± 75
I25.3	IL-8	CDR1mat	I1	432 ± 54
I25.6	IL-8	CDR1mat	I1	~ 750
I3.1	IL-8	CDR1mat	I1	~ 1000

K_Ds after affinity maturation (CDR1mat) are shown for different targets. Parental CDR3: CDR3 of binders after affinity maturation resulting from a clone of the initial library screen. n.d.: CDR3 was not found after sequencing of binders selected from initial screen.

5.5 Expression of selected vNAR domains

Previous work of Wittrup and colleagues showed that the dissociation constant calculated by affinity titration on the yeast surface correlates well with measurements of soluble protein.[146,189] To examine if this correlation is also consistent for vNAR domains selected by yeast surface display and also to assess thermal stability, we chose EpCAM binders H3 and H5, obtained from initial screenings with micromolar affinities as well as nanomolar EpCAM binder 5005, identified after CDR1 affinity maturation, for protein production in *E. coli* (**4.4**). The respective vNAR molecules were expressed in the pMX vector[183] that introduces maltose binding protein (MalE) and a hexahistine tag for fusion protein expression and purification. A scheme for protein production is depicted in **Figure 26** (**Fig. 26**). IgNAR V domains were expressed as MalE-fusion proteins with a triple alanine linker and cleavage site for tabacco etch virus protease (TEV) as well as a *C*-terminal triple alanine linker and hexa-histidine tag. vNARs were first purified using nickel affinity chromatography (IMAC), followed

by TEV-protease cleavage. The monomeric vNAR domains were further purified and isolated by gel filtration chromatography.

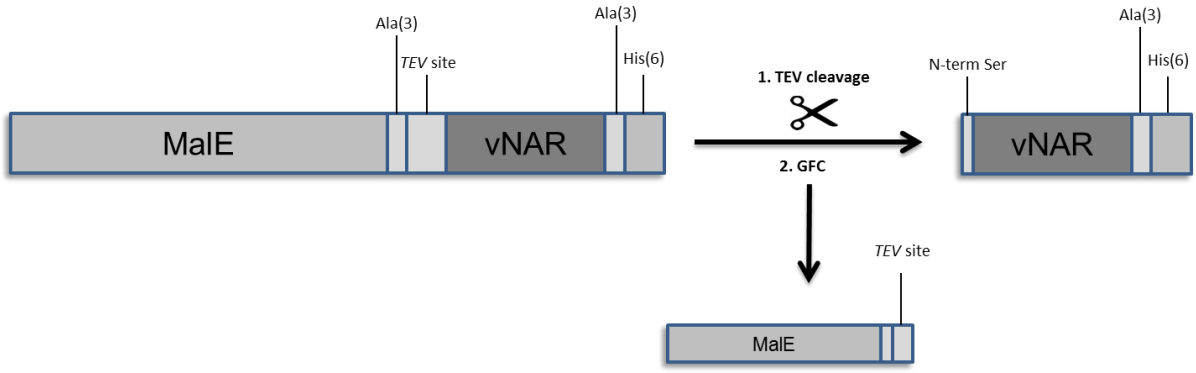


Figure 26. Scheme of of recombinant vNAR domain production as MalE-fusion protein. Ala(3): Triple alanine linker; His(6): hexahistidine-tag for protein purification; TEV site*:* site for specific cleavage by tabacco etch virus protease (TEV). IMAC purified protein was first digested *via* TEV protease and in a final step purified using GFC.

pMX vector, containing for *E.*coli codon-optimized α-EpCAM-vNAR DNA-sequences for variant 5005 was purchased from GeneArt (Life Technologies). Codon-optimization was performed according the manufacturer's software. For protein production (**4.4.1**) of αEpCAM-vNAR_H3 and α-EpCAM-vNAR_H5, the respective DNA fragment was PCR amplified using primer-combination pMX_vNAR_up/pMX_vNAR_lo and isolated total DNA from *S. cerevisiae* single clones H3 or H5, containing the plasmid DNA as template. Afterwards, PCR-product as well as pMX-vector was digested using *Acc* 65I and *Xba* I. After gel-purification, PCR-products were ligated into pMX vector. Electro-competent *E. coli* BMH 71-18 cells were transformed with plasmids pMX_αEpCAM-vNAR_H3, pMX_αEpCAM-vNAR_H5 and pMX_αEpCAM-vNAR_5005, respectively.

E. coli BMH 71-18 cells containing the respective plasmid were grown to an OD_{600} between 0.7 and 1 in 2 L dYT medium containing 25 mg/L chloramphenicol and protein production was induced with 1 mM IPTG. Cells were grown overnight (approx. 16 h) at 29 °C and harvested by centrifugation (**4.4.2**). The periplasmic fraction containing the respective MalE-vNAR fusion-protein was isolated. Respective recombinant vNAR proteins were purified by metal chelate affinity chromatography (HisTrap, GE Healthcare), equilibrated in PBS, pH 7.4. Proteins were eluted using a linear imidazole gradient starting from 0 % to 100 % of 1 M imidazole over 30 min (**Fig. 27**). The different vNAR proteins eluted at concentrations ranging from approximately 100 mM imidazole to 200 mM imidazole. MalE-vNAR fusion-proteins have a calculated molecular mass of approximately 57.5 kDa. As shown at **Figure 27** produced fusion-proteins migrate in the gel between 46 kDa and 58 kDa, correlating with the predicted molecular weight. Collected fractions shown on the gel were pooled for

each respective vNAR domain (**Fig. 27**). After this first purification step, fusion proteins were cleaved *via* tabacco etch virus protease cleavage overnight at 4 °C to isolate the solitary vNAR products.

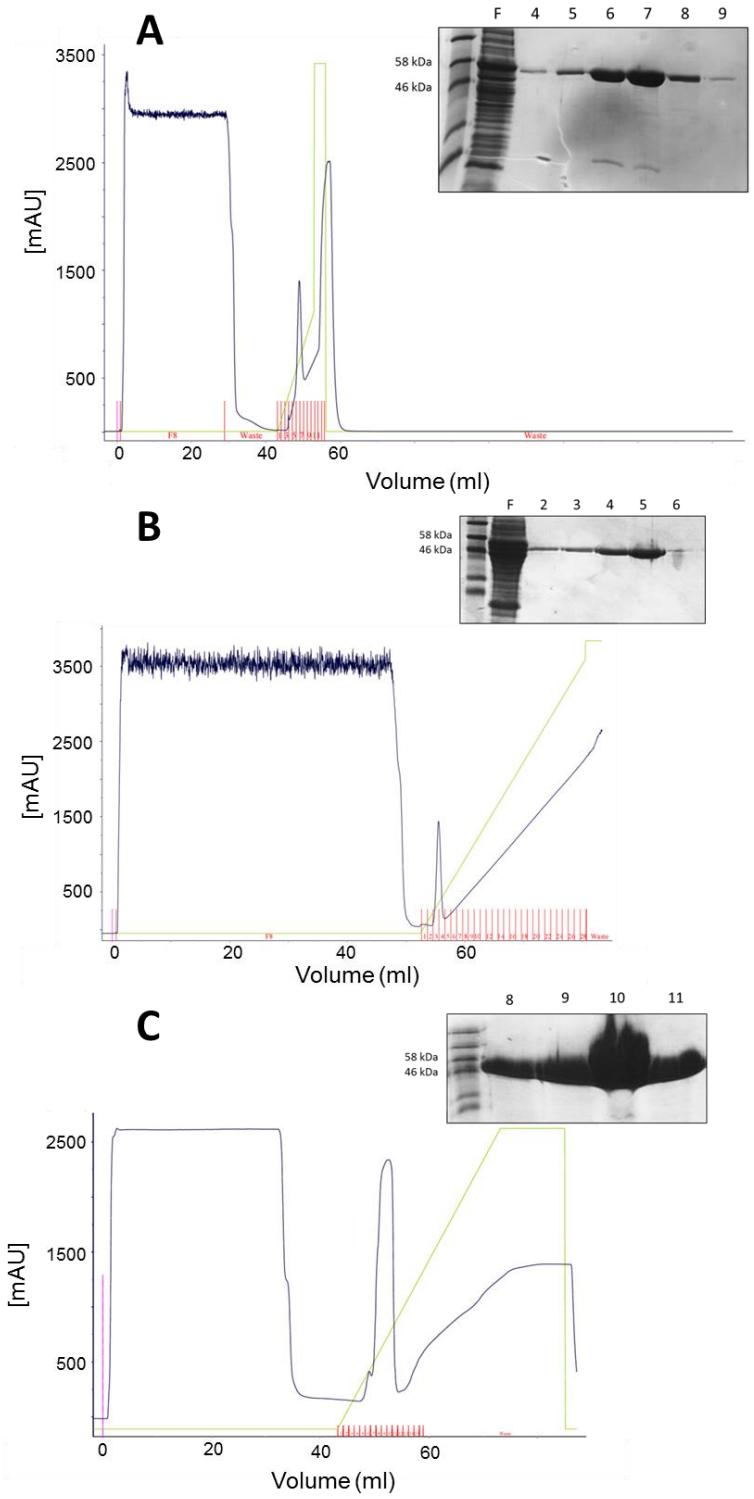


Figure 27. Elution profile of selected and recombinant produced MalE-vNAR fusion-proteins on a His Trap HP 1 ml column equilibrated in PBS, pH 7.4 and run with a flow rate of 1 ml/min. Blue: Absorption at 280 nM. Pink: Injection. Green: Imidazole gradient (0 % - 100 % of 1 M imidazole over 30 min). Red: Collected fractions. Corresponding gels of SDS-PAGE of collected fractions are shown. F: Flow-through. Numbers indicate collected fractions of the corresponding IMAC run. Chromatogram and gel (SDS-PAGE) of vNARs H3 (A), H5 (B) and 5005 (C).

Finally, proteins were purified to remove TEV-protease as well as MalE-protein using gel filtration on either a Superdex 200 pg 16/60 or on a Superdex 75 pg 16/60 (GE Healthcare), equilibrated in PBS, pH 7.4 (**Fig. 28**). The peak corresponding to the solitary vNAR domain was collected and concentrated using Amicon® Ultra-15 centrifugal filter devices (Merck Millipore Ltd.). Protein quality was assessed *via* SDS-PAGE indicating highly purified solitary vNAR proteins, migrating between 7 kDa and 17 kDa which is consistent with the predicted molecular mass of about 13.2 kDa.

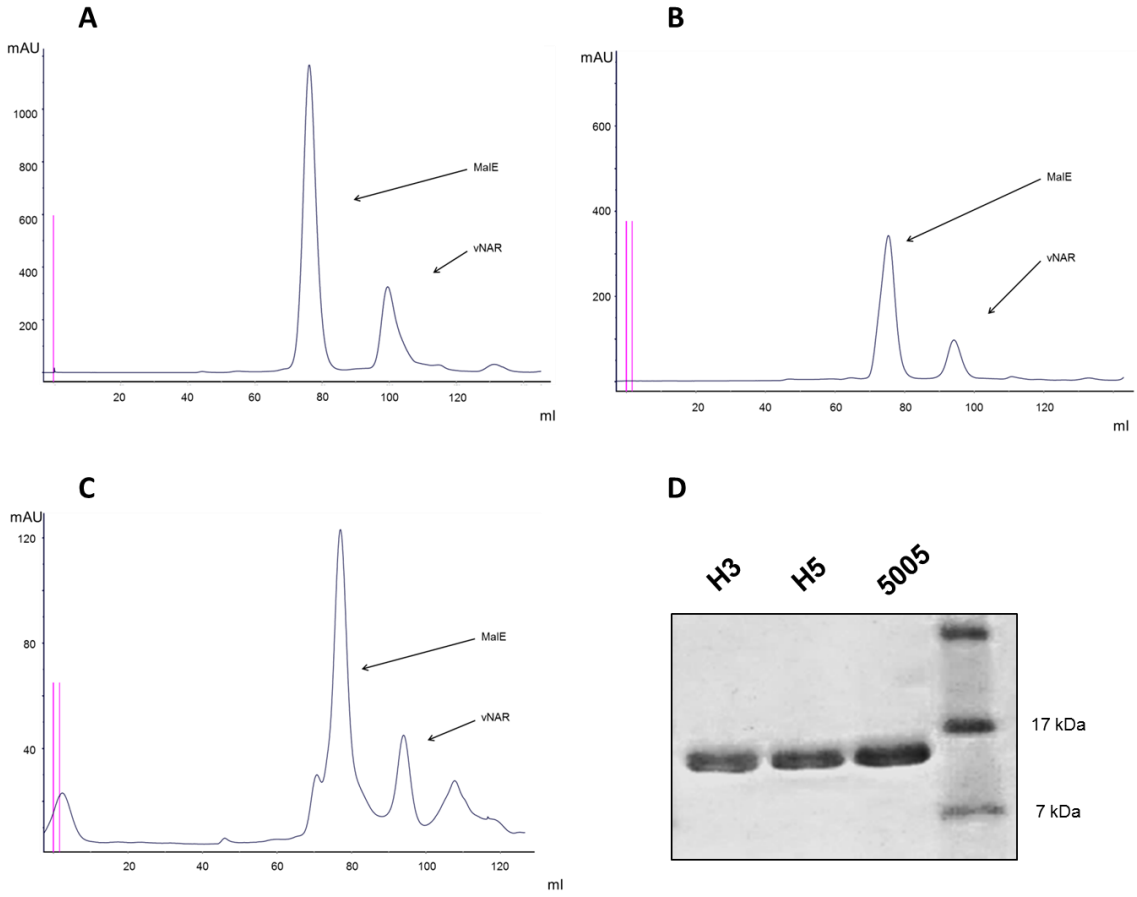


Figure 28. Purification of recombinant vNAR proteins *via* size exclusion chromatography (A-C) and quality control of purified vNAR products using SDS-PAGE (D). (A) Elution profile of IMAC-purified TEV-cleaved vNAR 5005 protein on a Superdex 200 pg 16/60 equilibrated in PBS, pH 7.4 and run with a flow rate of 1 ml/min. Elution profile of IMAC-purified TEV-cleaved vNAR H5 (B) and H3 (C) protein on a Superdex 75 pg 16/60 equilibrated in PBS, pH 7.4 and run with a flow rate of 1 ml/min. (D) Gel of SDS-PAGE of solitary vNAR domains following gel filtration.

Protein concentrations were determined on a photometer (Biospec-nano, Shimadzu), showing moderate yields for vNAR production as fusion-proteins following a two-step purification procedure with an additional manipulation step *via* TEV-protease cleavage. Although yields for α-EpCAM-vNAR 5005 were sufficient (approximately 1.1 mg/L), α-EpCAM-vNARs H3 and H5 were only produced in low amounts of about 0.3 mg/L and 0.6 mg/L, respectively.

5.6 Formatting of α-EpCAM-vNAR 5005 as Fc-fusion

To enhance vNAR expression, α-EpCAM-vNAR 5005 was genetically fused to the *N*-terminus of a human Fc domain (**4.5**). The rationale was that fusions to the human Fc-part, besides enhancing vNAR expression yields and causing the formation of dimers, can also elicit *in vivo* immune effector functions *via* ADCC and CDC. Moreover for cell-based immunofluorescence assays, binding of Fc-fusion proteins to cells e.g. tumor cells can easily be detected using fluorescently-labeled secondary antibodies targeting the human Fc-part.

For this, α-EpCAM-vNAR 5005 was expressed in the pExpress vector that introduces the human Fc-part. At first, α-EpCAM-vNAR 5005 was amplified using isolated pCT_αEpCAM-vNAR_5005 as template (**4.5.1**). The hinge-region was introduced in a consecutive two-step PCR with primer pair pEXPR_vNAR_up/pEXPR_vNAR_loI and pEXPR_vNAR_up/pEXPR_vNAR_loII, respectively. PCR product as well as pEXPR were digested using *Nhe* I and *Apa* I and the digested vNAR product was ligated into the pEXPR vector. Expi293F™ cells were transfected with pExpress_αEpCAM-vNAR_5005 using the Expi293™ Expression System Kit (Life Technologies) and cells were cultured at 37 °C, 5 % CO_2 (in air) and 100 rpm (**4.5.2**). After approximately 24 hours 150 µl of ExpiFectamine™ 293 transfection enhancer 1 and 1.5 ml ExpiFectamine™ 293 transfection enhancer 2 were added. Five days post transfection the suspension was centrifuged, cells were discarded and the supernatant containing the secreted vNAR-Fc fusion was harvested. Purification of vNAR-fusion protein was executed *via* Protein A affinity chromatography using a HiTrap Protein A HP 1 ml column (GE Healthcare), equilibrated in Protein A chromatography running buffer, pH 7 (**Fig. 29**). vNAR-Fc fusion protein was eluted using Protein A chromatography elution buffer, pH 3. To immediately neutralize acidic pH and thus, to minimize potential protein denaturation, 1 ml fractions were collected in tubes containing 200 µl of Protein A chromatography collection buffer, pH 9.0. Collected fractions were pooled and dialyzed against 5 L PBS, pH 7.4 overnight at 8 °C and protein quality was assessed *via* SDS-PAGE (**Fig. 29**). The vNAR-Fc fusion has a predicted molecular mass of about 43 kDa. The eluted protein migrates between 30 kDa and 46 kDa, which is consistent with the calculated molecular mass. The concentration of the protein was determined photometrically (Biospec-nano, Shimadzu) yielding in approximately 192 mg/L indicating high-level expression of vNAR-Fc fusion proteins in mammalian cells.

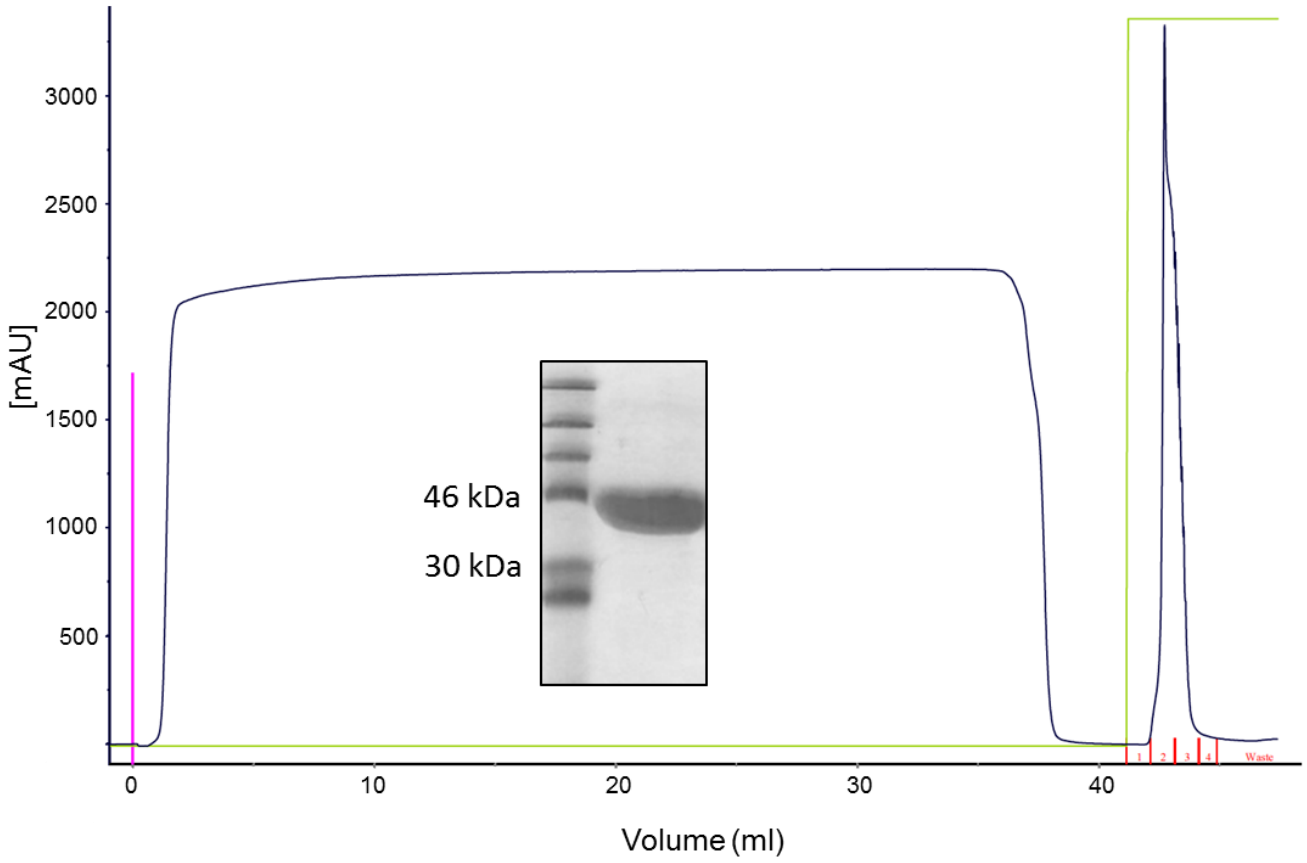


Figure 29. Elution profile of α-EpCAM_vNAR 5005-Fc fusion protein on a HiTrap Protein A HP 1 ml column equilibrated in Protein A chromatography running buffer, pH 7 and run with a flow rate of 1 ml/min. Blue: Absorption at 280 nM. Pink: Injection. Green: Protein A chromatography elution buffer concentration, pH 3 (%). Red: collected fractions. Corresponding gel of SDS-PAGE of pooled fractions is shown.

5.7 Characterization of selected vNAR domains

Soluble produced vNAR domains as well as formatted vNAR-Fc fusion protein were characterized in terms of affinity, selectivity and stability (**4.6**). Affinities of all produced vNARs were assessed using biolayer interferometry (FortéBio) and for the examination of thermal stability thermal shift assays were retrieved. Selectivity of α-EpCAM-vNAR 5005, its parental molecule α-EpCAM-vNAR H3 and of α-EpCAM-vNAR H5 were studied on the yeast surface. Off-target binding was analyzed with selected unrelated target proteins. Finally, formatted α-EpCAM-vNAR_5005-Fc fusion was used for tumor cell staining assays.

5.7.1 Affinity measurements using biolayer interferometry

The equilibrium dissociation constant (K_D) for all produced vNAR variants was determined on the Octet RED96 system (**4.6.1**). For solitary vNAR domains all assays were performed with Streptavidin Dip and Read™ biosensors in kinetics buffer. Sensors were loaded with biotinylated vNARs at

approximately 10 µg/ml. As negative control, a target-unspecific vNAR was used. Formatted α-EpCAM-vNAR 5005-Fc fusion was measured using Anti-Human IgG Fc Capture Dip and Read™ biosensors and the fusion protein was loaded at approximately 10 µg/ml (**Fig. 30**). Kinetic data sets were fitted using 1:1 Langmuir binding *via* the manufacturer's data analysis software with Savitzky-Golay filtering.

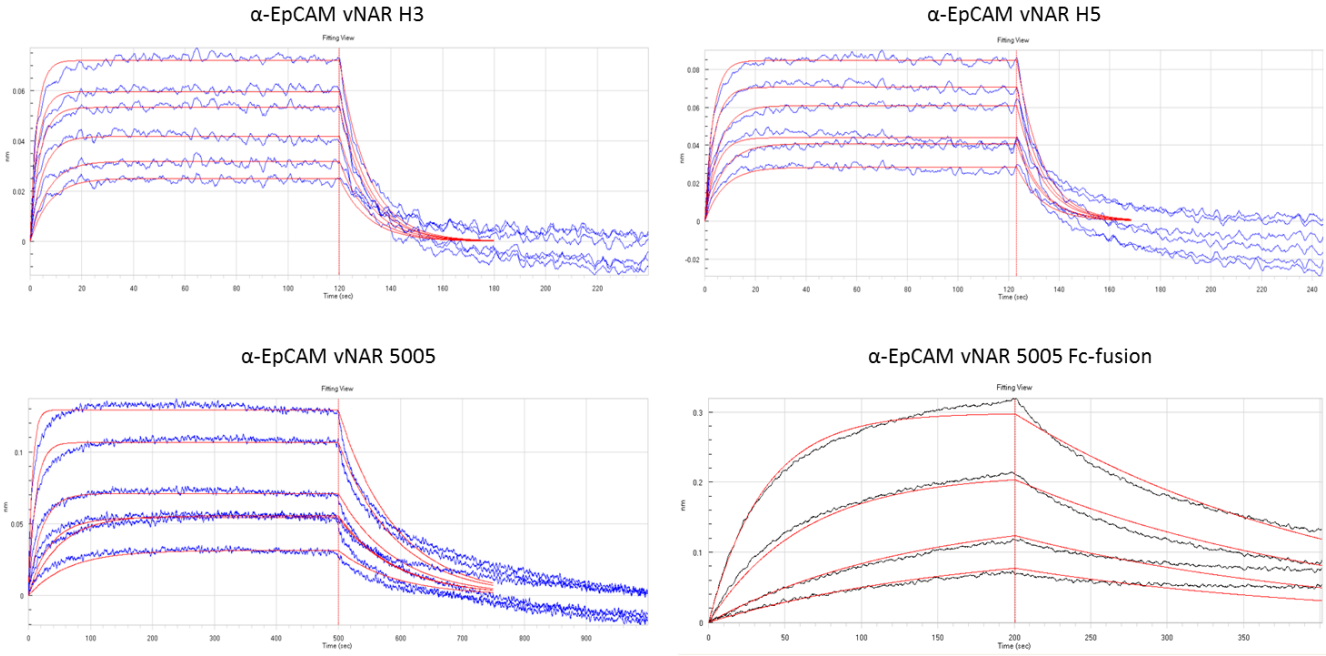


Figure 30. Binding curves of α-EpCAM vNARs used to determine K_D values *via* biolayer interferometry on the Octet® RED96 System. For all measurements, vNAR was immobilized on the tip surface. All vNARs except for the Fc-fusion were biotinylated. Correspondingly, Streptavidin Dip and Read™ biosensors were used. Anti-Human IgG Fc Capture Dip and Read™ biosensors were used for Fc-fusion measurements. For vNARs H3 and H5 measurements EpCAM concentrations of 5 µM, 4 µM, 3 µM, 2 µM, 1 µM and 0,75 µM were employed. Solitary vNAR 5005 was measured in 500 nM, 250 nM, 125 nM, 62.5 nM, 31.25 nM and 15.625 nM. Negative control with unrelated vNAR were tested at the highest respective concentration and subtracted from each curve. α-EpCAM vNAR 5005 Fc-fusion was measured in 67.5 nM, 33.8 nM, 8.4 nM, and 4.2 nM EpCAM. Negative control: kinetics buffer.

As shown in **Table 4**, dissociation constant calculated by affinity titration on the yeast surface correlates well with measurements of soluble protein (**Tab. 2**, **Tab. 3** and **Tab. 4**). Binder 5005 showed an approximately 40-fold enhanced affinity compared to the parental vNAR H3. Both, improvements in on-rate kinetics (5-fold) as well as in off-rate kinetics (8.1-fold) contributed to the enhanced dissociation constant. Formatting of vNARs as Fc fusion proteins does not negatively affect binding kinetics. Compared to the solitary vNAR 5005, we observed a marginal increased association. Additionally, there is evidence, that dimerization of vNAR 5005 mediated by the Fc part likely contributes to a significantly enhanced K_{off} compared to the solitary monomeric vNAR domain, causing a moderately improved equilibrium dissociation constant by the factor of about 3 (**Tab. 4**).

Table 4. Binding kinetics and Tm of soluble vNAR variants selected against EpCAM.

vNAR clone	Target	Screen	parental CDR3	K_{on} ($M^{-1}s^{-1}$)	K_{off} (s^{-1})	K_D [nM]	Tm [°C]
H3	EpCAM	initial	-	6.0×10^{4}	8.9×10^{-2}	1495	75.1 ± 0,5
H5	EpCAM	initial	-	4.3×10^{4}	1.0×10^{-1}	2342	71.5 ± 0,6
5005	EpCAM	CDR1mat	H3	3.0×10^{5}	1.1×10^{-2}	37	64.9 ± 0,6
5005-Fc	EpCAM	CDR1mat	H3	3.4×10^{5}	4.6×10^{-3}	13	-

Binding kinetics measured on the Octet® RED96 System. Tm determined *via* thermal shift assays.

5.7.2 Determination of thermal stability

The thermal stability of all three monomeric vNAR molecules was explored in thermal shift assays (**4.6.2**) on a BioRad 96CFX RT-PCR detection system using SYPRO Orange to obtain the melting temperature (**Fig. 31**, **Tab. 4**).Tm values varied from approximately 65 °C to over 75 °C, indicating a high thermal stability, with binder 5005 possessing the lowest Tm of 64.9 °C. However, its parental molecule, H3 showed the highest thermal stability (Tm = 75.1 °C), indicating that in this particular case enhancement of affinity is at the cost of thermal stability.

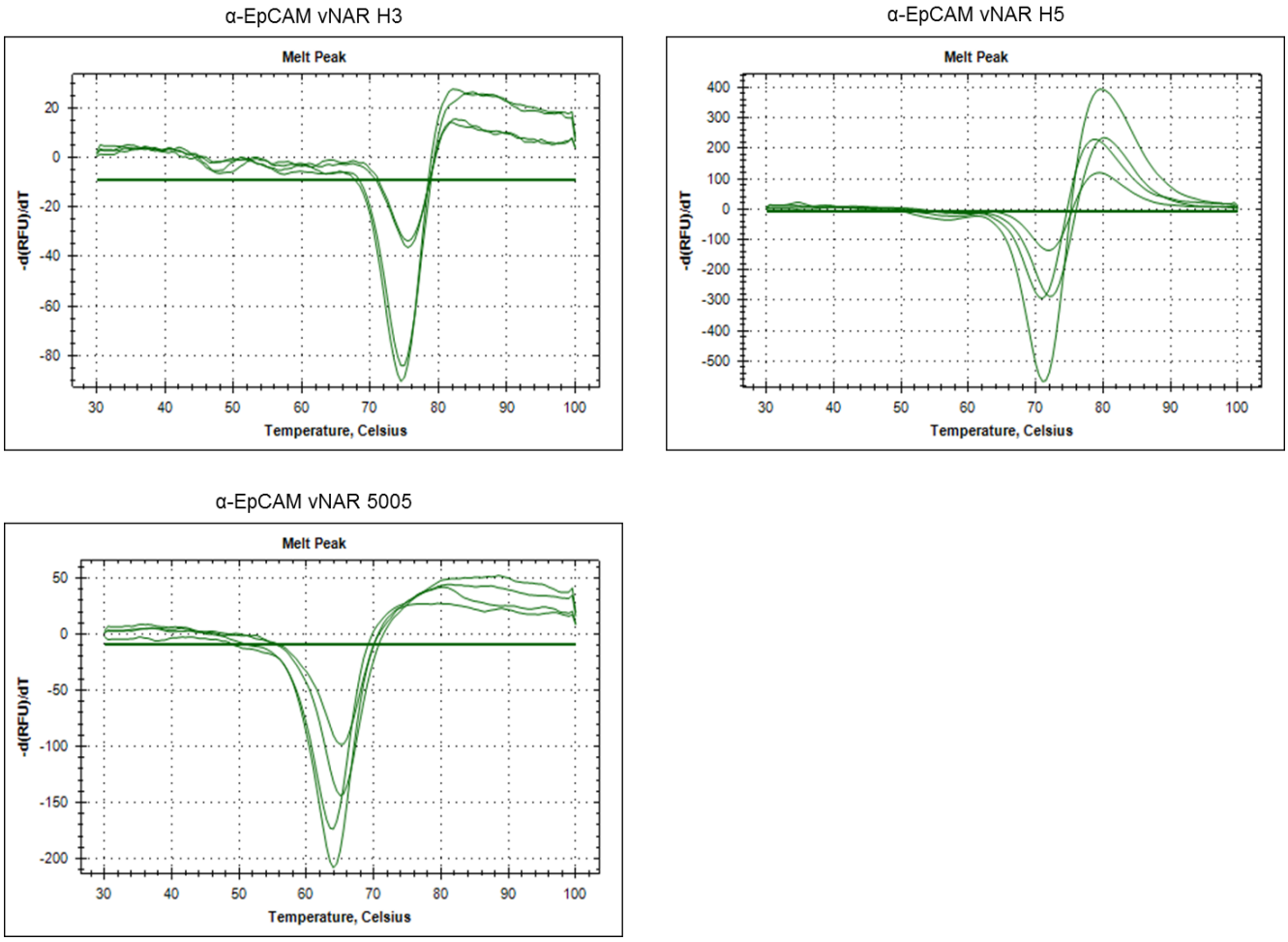


Figure 31. Thermal shift assays for Tm calculation of α-EpCAM vNARs. Melt peaks are shown for each individual vNAR. Quadruplicates were conducted to determine the standard deviation.

5.7.3 Determination of selectivity with related and unrelated target proteins using yeast surface display

Specificity of α-EpCAM-vNAR H3 and H5, isolated from initial screening as well as α-EpCAM-vNAR 5005 obtained after affinity maturation was scrutinized on the yeast surface (**4.3.2**) with three unrelated target-proteins, tGFP, murine EGFR and EphA2. Additionally, binding against murine EpCAM was assessed. For this, single clones were incubated with 1 µM of the respective related or unrelated target protein.

As shown in **Figure 32**, no off-target binding was detected against any of the selected unrelated target proteins, indicating high specificity of isolated, semi-synthetic vNAR domains (**Fig. 32**). However, all clones showed promiscuous binding to human and also murine EpCAM, an attribute which may be useful for animal studies.

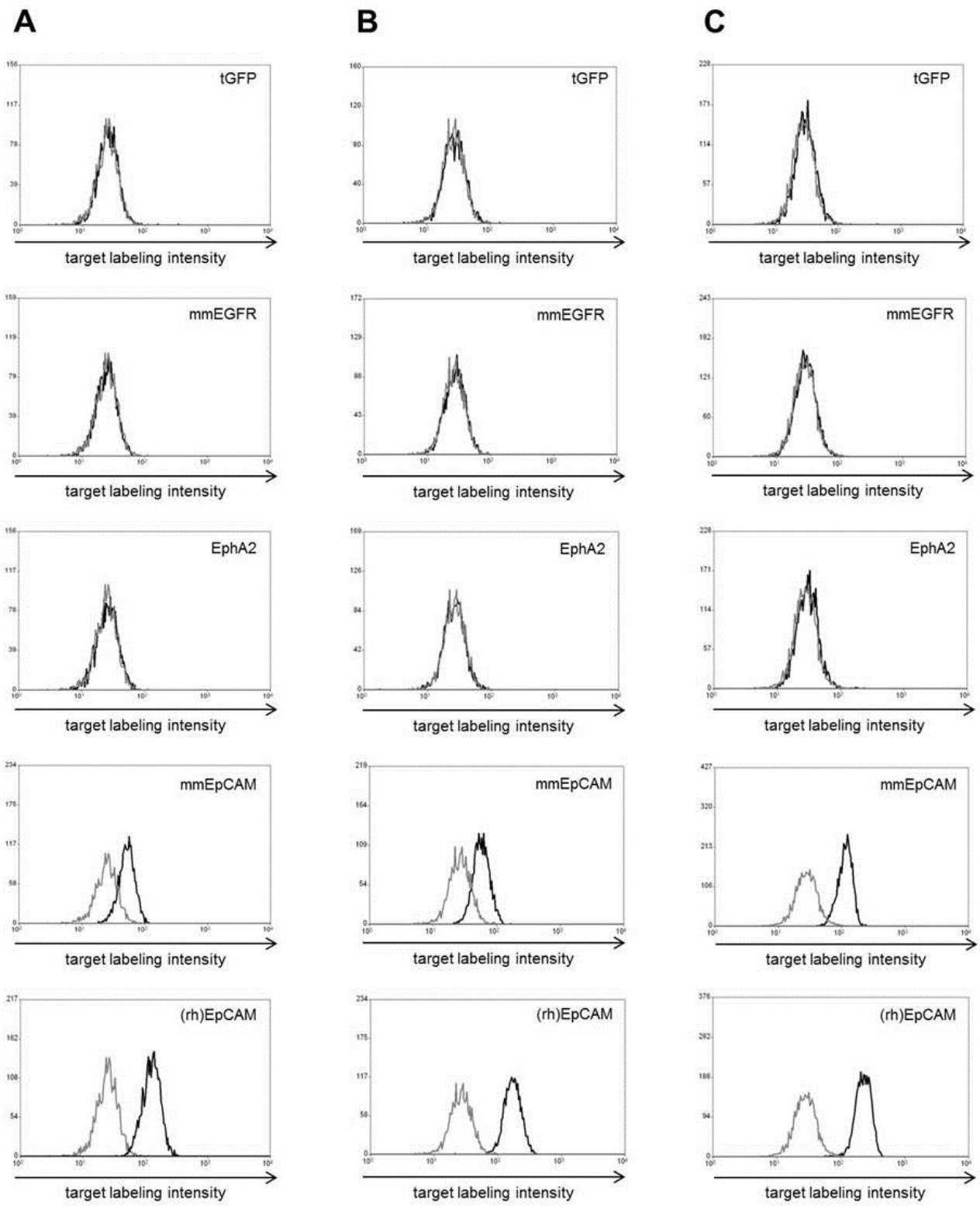


Figure 32. Specificity of α-EpCAM vNAR H3 (A), H5 (B) and 5005 (C) as determined *via* binding assays on the yeast surface. Off-target binding was validated against unrelated target proteins tGFP, murine EGFR (mm) and recombinant human EphA2. Target binding was verified against recombinant human (rh) EpCAM and murine (mm) EpCAM. cMyc normalized antigen binding is shown. Grey: negative control (cells stained for cMyc surface display and with Penta-His Alexa Fluor 488 conjugate). Black: target/off-target binding against 1 µM of the respective antigen (incubation time with target: approx. 30 min; 2nd staining-step with Penta-His Alexa Fluor 488 conjugate).

5.7.4 Tumor cell staining assays

To examine whether α-EpCAM-vNAR 5005 Fc-fusion protein can be used to selectively target EpCAM positive cells, tumor cell staining assays were conducted (**4.6.3**). In this study, four different cell lines were used, MCF-7, EBC-1(personal communication Achim Doerner, Merck Serono Darmstadt) and T47D, considered as EpCAM positive cell lines as well as CHO(s) cells, generally recognized as EpCAM negative.[190-194]

For this purpose, analysis of cell binding was performed by FACS analysis. Approximately 5 x 10^5 cells of the respective cell line were centrifuged and used for each antibody and vNAR staining. As positive control, EpCAM-specific antibody clone HEA-125 conjugated to PE was used in a dilution of 1:11, according to the manufacturer's instructions (MACS, Miltenyi Biotec). Cells were stained for approximately 1 hour at 4 °C with α-EpCAM-vNAR 5005 Fc fusion protein in different concentrations or with the control antibody. After washing and incubation with anti-human IgG conjugated to PE (Fcγ-specific, eBioscience), except for EpCAM-specific antibody HEA-125, cells were analyzed.

As shown in **Figure 33**, the vNAR-Fc fusion bound significantly to all EpCAM positive cell lines (**Fig. 33**). However, strongest binding was observed for CHO cells, whereas the positive control, EpCAM-specific clone HEA-125 did not show any binding, giving clear evidence for strong off-target binding in a concentration-dependent manner. Hence, it was unfeasible to use α-EpCAM-vNAR 5005 for *in vivo* studies such as tumor *xenograft* models.

This observation had several consequences for the follow-up project which was the construction of a vNAR domain, simultaneously targeting two different antigens. For this engineering approach, vNAR 5005 was used as starting molecule, because at the time this particular project was started, evidence for unspecific binding of this antibody domain was missing. Due to the unfeasibility of the utilization of a bi-specific molecule based on vNAR 5005 for cell assays as well as *in vivo* studies, bi-specificity of engineered vNAR domains could only be validated using different methodologies of *in vitro* characterization.

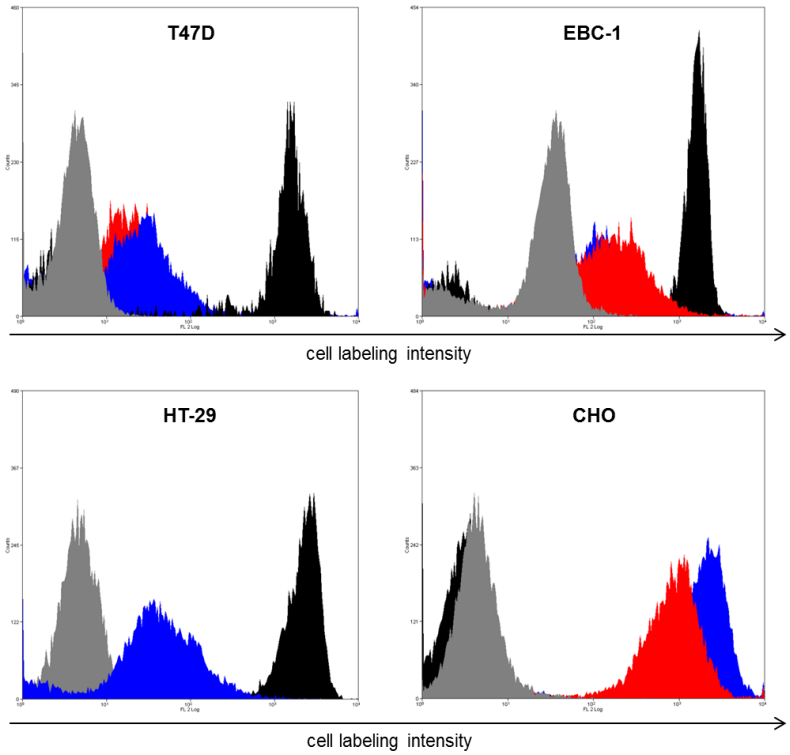


Figure 33. Specificity of α-EpCAM vNAR 5005 Fc fusion protein validated by tumor cell staining assays. Different cell-lines are shown. Cells were incubated with vNAR Fc fusion protein or EpCAM-specific control antibody for approx. 1 h, washed and incubated with Fc-specific secondary antibody-PE conjugate (except for control antibody). Grey: Cells labeled with Fc-specific secondary antibody-PE conjugate only. Black: Cells labeled with EpCAM-specific control antibody. Blue: Cells labeled with 1 µM α-EpCAM vNAR 5005 Fc fusion protein. Red: Cells labeled with 0.5 µM α-EpCAM vNAR 5005 Fc fusion protein.

5.8 Generation of a new paratope into the vNAR scaffold through randomization of HV2

The following chapter describes the generation of a new antigen-binding site into HV2 of the vNAR domain. Through the randomization of this surface-exposed loop and library screening, binders were isolated with the capability to target distinct antigens with HV2. Hence, HV2 was used to engineer a second paratope that acts independently from the conventional antigen-binding site, composed of CDR3 and CDR1 (and potentially HV4). Importantly, as validated using yeast surface display, the establishment of this new and unique feature did not impair the conventional paratope's affinity for its target, clearly indicating the structural integrity of the vNAR domain. The rationale for this was that for vNAR type II and type IV, HV2 is located distantly from the conventional paratope (**Fig. 34**). Note that crystallography revealed that for type I vNAR domains, CDR3 is located in close proximity to HV2.[24] It was shown that HV2 is prone to mutations [98], accordingly, we believed that the vNAR overall structure tolerates mutations in this region without largely compromising its structural integrity. Therefore, α-EpCAM-vNAR 5005 was used as starting molecule, which targets EpCAM

with CDR3 and CDR1. Through the diversification of HV2, we wanted to select for binders targeting a new antigen with HV2 whilst retaining its ability to target EpCAM.

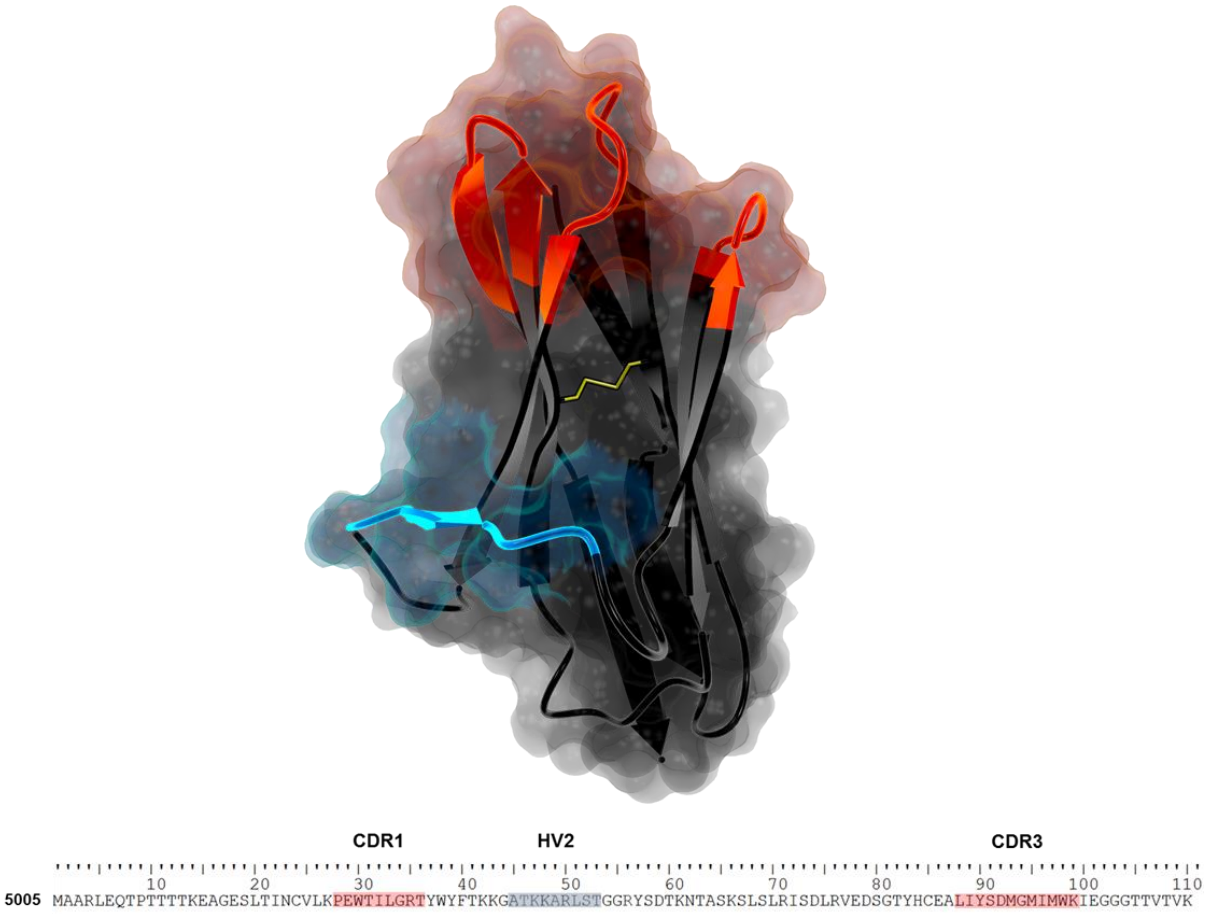


Figure 34. Depiction of the rationale for the generation of a new antigen-binding site into the vNAR scaffold. The conventional paratope composed of CDR3, CDR1 (and HV4) is shown in red. The potentially new antigen-binding site, consisting of HV2 is shown in blue. Yellow: Disulfide bond. Model based on pdb entry 4HGK [103] generated using YASARA structure.[107] Sequence shown for parental molecule 5005 used for library design. Sequence for CDR1 and CDR3 as well as HV4 of the conventional antigen-binding site shaded in red. Residues in sequence exposed loop corresponding to HV2 and considered for randomization shown in blue. Figure was constructed in close cooperation with Martin Empting.

5.8.1 Construction of a HV2 randomized library based on EpCAM-specific vNAR 5005

For the generation of a HV2 randomized library based on EpCAM-specific vNAR 5005, the pCT plasmid, isolated from the respective single clone was used as starting material (**4.2.3**). Akin to randomization of CDR3 and CDR1, diversifications were performed by incorporation of triplet codons into the corresponding oligonucleotide. Nine residues of the surface-exposed loop HV2 were chosen for randomization. Library was established in a consecutive two-step splicing by overlap extension PCR, as schematically depicted in **Figure 35** (**Fig. 35**). For the first PCR step, two reactions were carried out in parallel. On the one hand, primer pair HV2_SOE_rand_up/pCT_Seq_lo was used to

randomize nine residues within HV2. pCT_Seq_lo hybridized at the plasmid backbone, to conserve CDR3 as well as framework region 4. On the other hand, to generate the corresponding low vNAR fragment, primer pair pCT_Seq_up/HV2_SOE_lo was employed. Afterwards, 1 µl of PCR-products were used as template for the follow-up second PCR reaction, respectively. After six cycles, primer pair pCT_Seq_up/pCT_Seq_lo was added to generate the complete HV2-randomized vNAR fragment with sites for homologous recombination in *S. cerevisiae*.

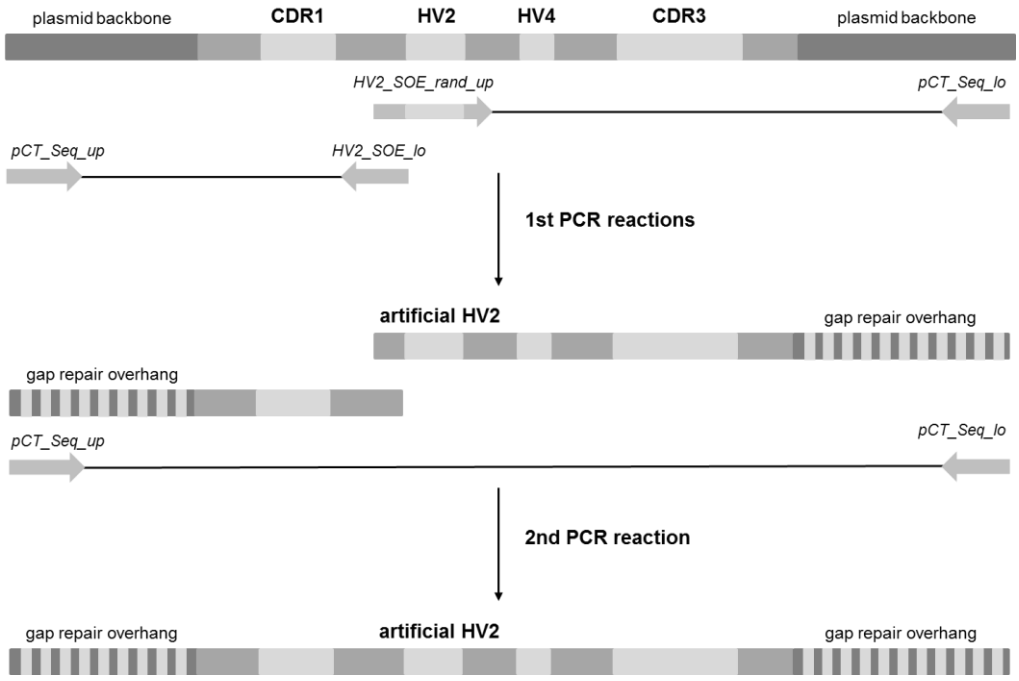


Figure 35. Schematic representation of PCR-based randomization of HV2. pCT plasmid encoding EpCAM-specific vNAR 5005 was used as template. In first PCRs HV2 was randomized and the corresponding up and low fragments of the vNAR domain were constructed. In a subsequent reaction the total vNAR molecule was constructed *via* splicing by overlap extension.

A yeast surface display library was established akin to initial library construction, as described in **chapter 5.2**. This yielded in a library diversity of approximately 10^9 unique clones as determined by dilution plating.

Library quality was assessed using yeast surface display with regard to cMyc-labeling (**4.3.1**) and EpCAM-labeling (**4.3.2**, **Fig. 36**). Within one day post-induction, there was approximately 40 % of cMyc-labeling, giving evidence for expression and high copy number display of HV2-diversified vNAR domains on the surface of *S. cerevisiae*. Moreover, we were able to detect about 36 % of EpCAM-binding, indicating that the vast majority of displayed vNARs, which are progenies of α-EpCAM-vNAR 5005, retain their structural and functional integrity.

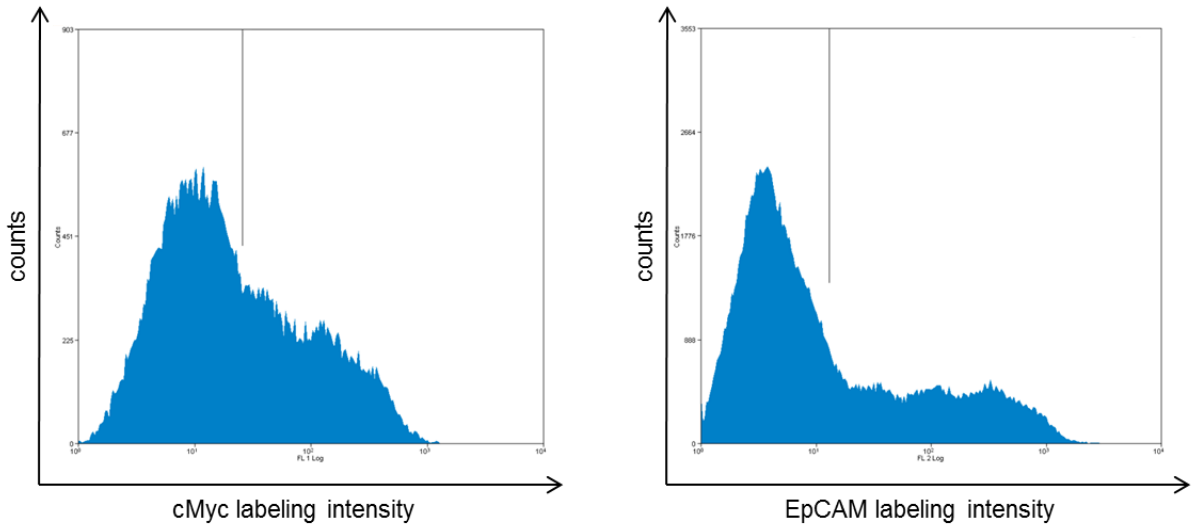


Figure 36. Analysis of HV2-randomized library based on EpCAM-specific vNAR 5005. Histogram of cMyc surface expression (left) and EpCAM binding (1 µM EpCAM, right) of the constructed vNAR library assessed by indirect immunofluorescence labeling and flow cytometry one day post induction. Cells in gate: 40.1 % for cMyc staining and 35.9 % for EpCAM-labeling (detected with EpCAM-specific antibody clone HEA-125 conjugated to PE).

5.8.2 Selection of vNAR molecules targeting CD3ε

Recombinant human cluster of differentiation 3 subunit ε (CD3ε) was purchased from AcroBiosystems. It consists of amino acids Asp23-Asp126 fused to a hexa-histidine tag for detection. The calculated molecular weight is predicted to be 13.2 kDa.

To isolate single clones, that bind CD3ε while retaining high affinity for EpCAM, the first round of screening was performed two-dimensionally for target binding, using his-tagged EpCAM and biotinylated CD3ε (**4.3.4**). Hence, antigen binding was detected using Alexa Fluor 488 labeled anti-His-tag antibody as well as streptavidin-APC. In the first round of screening, approximately 2 x 10^8 clones were sorted (**Fig. 37**). In the second and third round, cells were only sorted for CD3ε-binding, since cMyc surface labeling was around 50 %. To avoid off-target binding against detection reagents, the labeling strategy was alternated from using biotinylated CD3ε to his-tagged antigen which was detected *via* Alexa Fluor 488 labeled anti-His-tag antibody. Moreover, the target concentration was reduced in the third screening round to enhance stringency. For this reason and also due to the presence of a minor target-positive population, a resort was performed after round three. The fourth round of sorting revealed a significant enrichment of a double target-positive population. Peculiarly, as shown in **Figure 37**, also only CD3ε-positive clones were enriched. However, since we were only interested in bi-specific clones, we exclusively selected for EpCAM- and CD3ε-positive populations (**Fig. 37**).

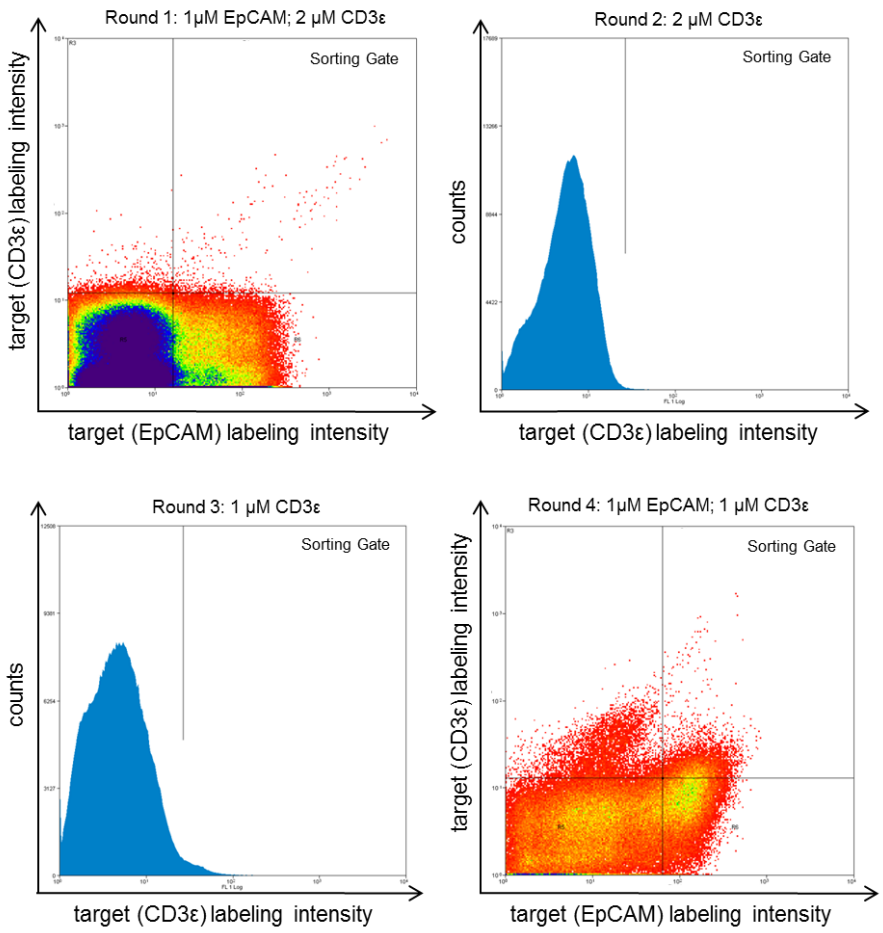


Figure 37. Screening of a HV2-randomized EpCAM-specific library based on vNAR 5005 against CD3ε. Sorting gates and target concentrations are shown. In round one and four, cells were simultaneously labeled for EpCAM binding using Penta-His Alexa Fluor 488 conjugate and CD3ε binding *via* biotinylated antigen and Steptavidin-APC. In round two and three, cells were only labeled for CD3ε binding using his-tagged antigen and Penta-His Alexa Fluor 488 conjugate. After round 3 a resort was performed. Cells in the sorting gate were isolated, grown and induced for the next round of selection. Target-positive cells in sorting gate: (R1) 0.1 %; (R2) 0.17 %; (R3) 1.92 % and (R4) 5.8 %.

Single clones were analyzed in terms of their ability to target both antigens (**4.3.2**). As shown for clones αEpCAM-CD3ε-vNAR H5 and αEpCAM-CD3ε-vNAR B1 in **Figure 38**, several double-positive clones were identified, whereas for the parental molecule α-EpCAM- vNAR 5005 no binding was observed against CD3ε (**Fig. 40**). Most of the clones still bound strongly to EpCAM at a concentration of 1 µM, but binding to CD3ε was very weak at 1 µM, as exemplarily shown for single clone αEpCAM-CD3ε-vNAR H5. Notwithstanding, for one clone, referred to as αEpCAM-CD3ε-vNAR B1 we observed binding to CD3ε to a significantly higher extent compared to all other analyzed clones. However, binding to EpCAM was only moderate, indicating that in this particular case the randomized HV2 sequence impairs EpCAM-binding at the conventional antigen-binding site. Sequencing of double positive clones revealed several different, potentially bi-specific vNARs (**Fig. 39**). Since all but one single clone only showed very weak binding to CD3ε, we decided only to characterize αEpCAM-CD3ε-vNAR B1 in more detail.

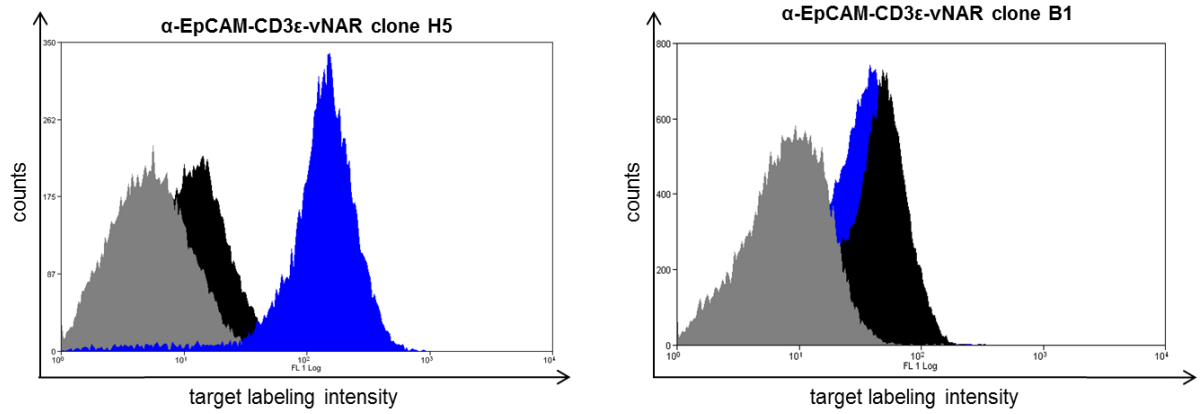


Figure 38. Single clone analysis of potentially bi-specific vNARs after randomization of HV2 and library screening for EpCAM and CD3ε binding. Single clones were incubated with 1 µM of the respective antigen. Alexa Fluor 488 labeled anti-His-tag antibody was used for the detection of target binding. Blue: Cells stained with EpCAM. Black: Cells incubated with CD3ε. Grey: Negative control, cells only labeled with secondary detection antibody.

Affinities of the potentially bi-specific clone α-EpCAM-CD3ε-vNAR B1 as well as its parental clone α-EpCAM-vNAR 5005 were assessed *via* affinity titration on the yeast surface (**4.3.3**, **Tab. 5**). α-EpCAM-CD3ε-vNAR B1 showed an affinity for CD3ε of about 409 nM. However, the affinity against EpCAM, declined substantially compared to the parental clone. As shown in **Figure 39**, bi-specific clone α-EpCAM-CD3ε-vNAR B1 bears an Arg56Gln mutation at the interface between the surface-exposed loop corresponding to HV2 and the adjacent framework region, referred to as framework region 3a (Fw3a). Importantly, this residue was not considered for library design.

Table 5. Equilibrium dissociation constants (K_D) determined by yeast surface display for bi-specific HV-randomized clones compared to the parental clone.

vNAR clone	**Type of molecule**	**K_D [nM] EpCAM**	**K_D [nM] CD3ε/Fcγ**
5005	Parental molecule	40 ± 13	- (CD3ε/ Fcγ)
B1	After library screen (Gly-**Gln**-Tyr motif) **CD3ε**	~ 1500	409 ± 70 (CD3ε)
B1mut	Q → R framework (Gly-**Arg**-Tyr motif) **CD3ε**	46 ± 8	422 ± 60 (CD3ε)
F1	After library screen human **Fcγ**	53 ± 10	~ 2600 (Fcγ)

Sequence comparison of different vNAR types and also of different shark species revealed a conserved Gly-Arg-Tyr motif (Gly55 – Tyr57) at the interface between HV2 and Fw3a (**Appendix B**). To determine whether the mutated Gly-Gln-Tyr motif has a negative impact on EpCAM-binding for bi-specific clone B1, we employed a backward mutation (Gly-**Gln**-Tyr to Gly-**Arg**-Tyr) *via* SOE-PCR (**4.2.3**). The PCR product was established in a consecutive two-step PCR, akin to HV2-randomization (**Fig. 31**). For the first PCR step, two reactions were carried out in parallel. On the one hand, primer pair HV2_B1mut_up/pCT_Seq_lo was used to reincorporate the Arg residue at the desired position. On the other hand, to generate the corresponding low vNAR fragment, primer pair pCT_Seq_up/ HV2_B1mut_lo was used. Afterwards, 1 µl of PCR-products were used as template for the follow-up second PCR reaction, respectively. After six cycles, primer pair pCT_Seq_up/pCT_Seq_lo was added to generate the complete mutated vNAR B1mut fragment with sites for homologous recombination in *S. cerevisiae*. EBY 100 cells were transformed with the described PCR product and with pCT plasmid.

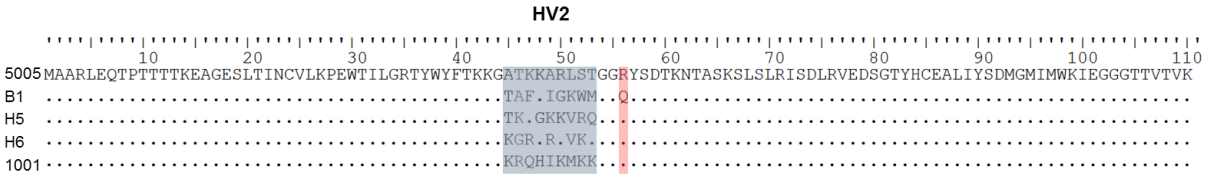


Figure 39. Sequences of potentially bi-specific EpCAM binding and CD3ε binding vNARs aligned to parental mono-specific molecule 5005. Dots indicate identical residues. Blue: Residues in surface-exposed loop corresponding to HV2 considered for library design. Red: Residue at the interface of HV2 and FW3a not considered for library design.

EpCAM-binding and binding to CD3ε was evaluated on the yeast surface using 1 µM of target (**4.3.2**), respectively, and compared to the parental molecule 5005 (**Fig. 40**). Equilibrium dissociation constant titration (**4.3.3**) confirmed the observation, that affinity for EpCAM was restored through the backward mutation to the original framework (Gly-Arg-Tyr motif), that is conserved amongst shark species as well as vNAR types (**Tab. 5**). Importantly, conversion into the original framework did not significantly influence the affinity for CD3ε. Essentially, these findings are giving clear evidence that bi-specific vNAR molecules can be generated using this novel approach in which HV2 facilitates antigen binding.

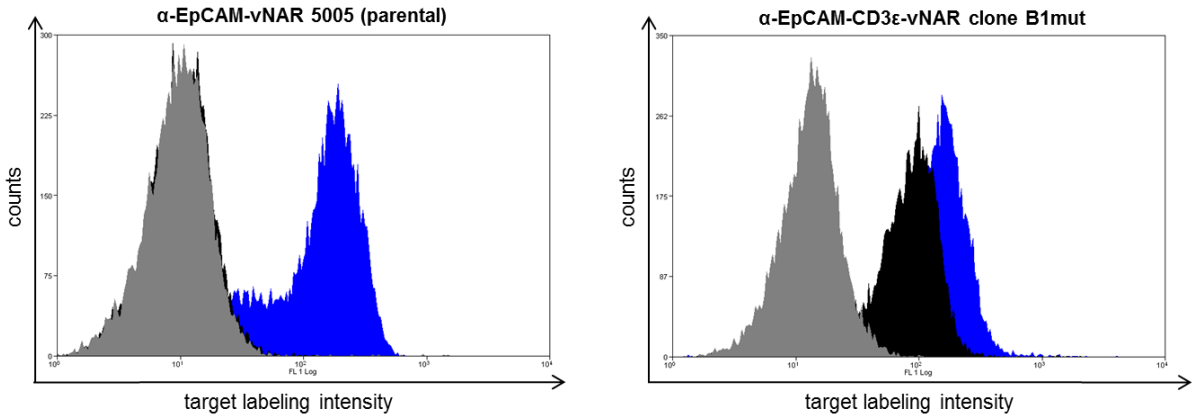


Figure 40. Single clone analysis of mono-specific, parental clone 5005 and HV2-diversified clone B1 for EpCAM and CD3ε binding. Single clones were incubated with 1 µM of the respective antigen. Alexa Fluor 488 labeled anti-His-tag antibody was used for the detection of target binding. Blue: Cells stained with EpCAM. Black: Cells incubated with CD3ε. Grey: Negative control, cells only labeled with secondary detection antibody.

5.8.3 Selection of vNAR molecules targeting Fcγ

In order to study whether this new strategy for the generation of a new antigen-binding site into the vNAR scaffold is generally applicable, the same library was also screened to select for binders, simultaneously addressing EpCAM and human Fcγ. For this, a non-glycosylated IL8-vNAR-Fcγ fusion protein (made in-house) was used as target protein.

In the first round of sorting approximately 2 x 10^8 cells were screened (**4.3.4**, **Fig. 41**). Binding to EpCAM was detected using Alexa Fluor 488 labeled anti-His-tag antibody and binding to Fcγ was detected using anti-human IgG conjugated to PE (Fcγ-specific, eBioscience). To avoid off-target binding against detection reagents, the second and third round of sorting was carried out only one-dimensionally. In the second round of screening, Fcγ-binding was detected *via* a Fcγ-specific Fab-fragment conjugated to Alexa Fluor 488. In the third round of sorting this secondary detection reagent was once again alternated to anti-human IgG PE-conjugate. In this round, a significant enrichment of target-positive populations was detected. To isolate binders that target both, EpCAM and human Fcγ, in the fourth round, sorting was once again switched to two-dimensionally screening for simultaneous detection of bi-specific target binding.

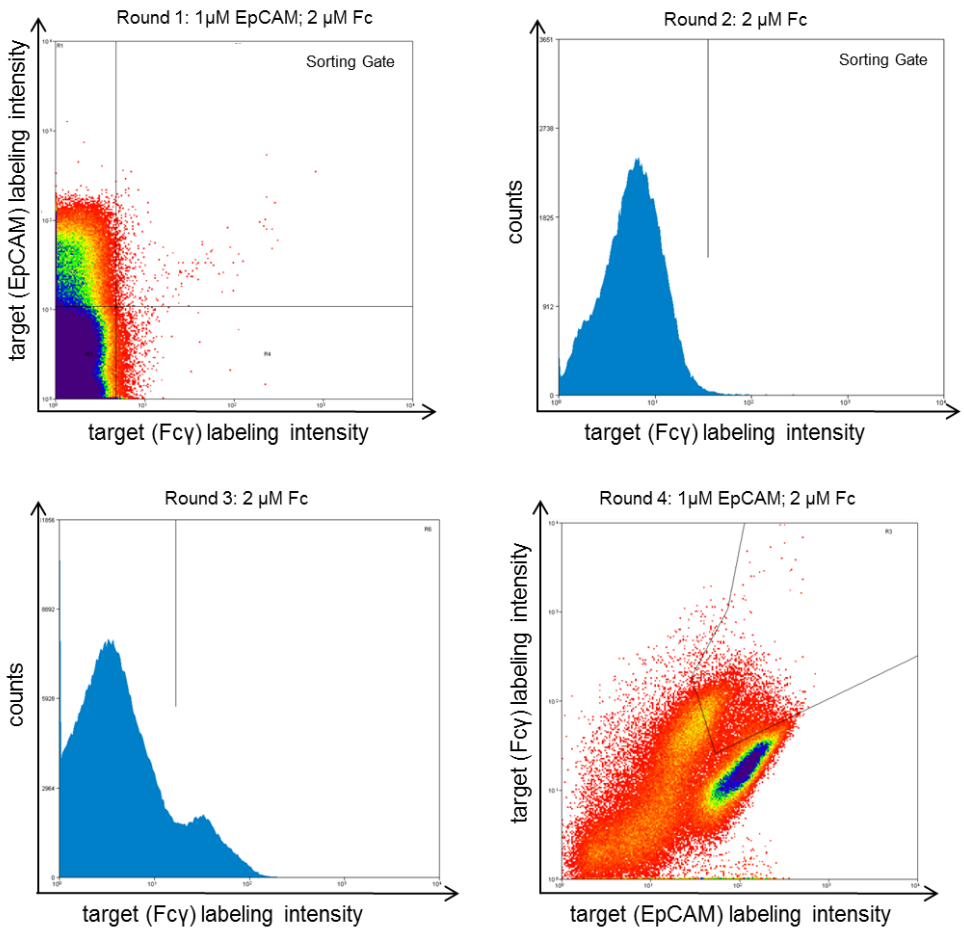


Figure 41. Screening of a HV2-randomized EpCAM-specific library for bi-specificity against human Fcγ. Sorting gates and target concentrations are shown. In round one and four, cells were simultaneously labeled for EpCAM- and Fcγ-binding. In round two and three, cells were only labeled for Fcγ-binding. Cells in the sorting gate were isolated, grown and induced for the next round of selection. Target-positive cells after sorting gate: (R1) 0.11 %; (R2) 0.41 %; (R3) 14.68 % and (R4) 9.21 %.

Single clones were analyzed for bi-specificity (**4.3.2**) and double-positive clones were sent out for sequencing, resulting in only one single sequence (**Fig. 42**). Characterization of this clone (**4.3.3**), referred to as α-EpCAM-Fcγ-vNAR F1, resulted in a low affinity against human Fcγ of approximately 2.6 µM without significantly compromising the affinity against EpCAM compared to its parental clone 5005, as listed in **Table 5** (**Tab. 5**). Of note, parental molecule 5005 does not display any affinity for human Fcγ. Interestingly, clone F1 only bound to non-glycosylated human Fcγ, as investigated with solitary Fc-fragments (provided by Stephan Dickgießer, Technische Universität Darmstadt). For the glycosylated equivalent we were unable to detect any binding at the highest concentrations tested (**Fig. 43**). Hence, one can draw conclusions with regard to the epitope that is addressed by α-EpCAM-Fcγ-vNAR F1. It is tempting to speculate, that position 297 of Fcγ is involved in the vNAR-target-interaction, since for wild-type Fcγ, this residue is Asn which is typically glycosylated. However, the provided unglycosylated Fcγ contains an Asn to Ala mutation at this position and it can

be assumed that this mutation, that prevents glycosylation, is part of the epitope targeted by α-EpCAM-Fcγ-vNAR F1. Another hypothesis is that glycosylation of the Fc-fragment covers the epitope targeted by α-EpCAM-Fcγ-vNAR F1 in a way that this region as structurally not accessible.

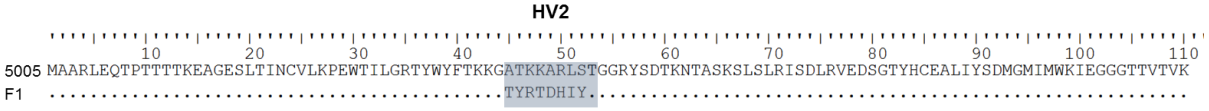


Figure 42. Sequences of bi-specific EpCAM binding and Fcγ binding vNAR F1 aligned to parental mono-specific molecule 5005. Dots indicate identical residues. Blue: Residues in surface-exposed loop corresponding to HV2 considered for library design

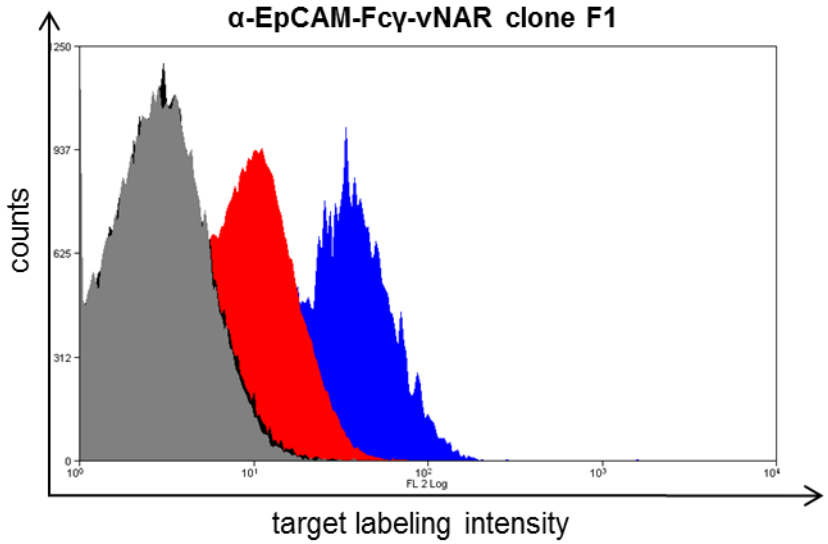


Figure 43. Single clone analysis of EpCAM and glycosylated Fcγ binding bi-specific clone F1. Single clone was incubated with 1 µM of the respective antigen. Anti-human IgG conjugated to PE (Fcγ-specific, eBioscience) was used for the detection of target binding. Grey: Negative control, cells only labeled with secondary detection antibody. Black: Cells incubated with glycosylated Fcγ. Red: Cells labeled with non-glycosylated Fcγ. Blue: Cells labeled with EpCAM and detected *via* EpCAM-specific antibody clone HEA-125 conjugated to PE.

6. DISCUSSION

6.1 Library construction, selection of antigen-specific vNAR molecules and affinity maturation of enriched target binding variants

Within the scope of this work we successfully developed an *in vitro* affinity maturation process for the generation of high-affinity vNAR domains using yeast surface display as platform technology. The isolation of affinity-matured binders was performed in a two-step process. Screening from a semi-synthetic vNAR library in which CDR3 was totally randomized resulted in the selection of binders comprising moderate affinities to their target, that were improved by diversification of CDR1 and sublibrary screening.

Initially, binders were selected against EpCAM, EphA2, HTRA1[187] and IL-8, clearly demonstrating that vNAR molecules can be isolated from this semi-synthetic library to target a wide range of different antigens. Affinity maturation was employed for EpCAM, HTRA1,[187] and IL-8 binding vNAR domains, giving evidence, that this new methodology is generally applicable. Additionally, a subset of binders was produced as soluble proteins in *E. coli* and *in vitro* characterization using biolayer interferometry confirmed the affinity maturation process and affinities as determined on the yeast surface.

Deep sequencing of natural bamboo shark vNARs from non-immunized animals confirmed that the primary vNAR diversity of this species is mainly restricted to CDR3, while the other regions (CDR1, HV2 and HV4) demonstrate only minor variations.[109] Expanding the set of variants by generating a semi-synthetic library of 2×10^8 variants with fully randomized CDR3 proved to be a valid strategy to obtain antigen binding vNAR molecules. In this respect, previous experiments failed to isolate EpCAM-specific vNAR molecules from the naïve repertoire of *Chiloscyllium plagiosum* (not shown). Regarding this, it is worthy to mention, that the total number of unique clones in this library was comparable to those of the herein presented semi-synthetic library.

Interestingly, in the original vNAR repertoire of the bamboo shark, typically one cysteine residue in CDR1 was conserved that presumably forms a disulfide bond with a Cys residue in CDR3 (**Fig. 1C**). In sharks three different types of vNARs have been characterized based on the position of non-canonical Cys residues.[51] An additional type of IgNAR, termed type IV, possesses only one canonical disulfide bond. Except IgNAR type III, which has limited sequence diversity in CDR3, all types give rise to antibodies with high affinities.[100,101,104,106,113,115] The semisynthetic library is cysteine-free in CDR3 and also has Cys in CDR1 replaced by Tyr. Therefore, it may be possible that the enhanced conformational flexibility of CDR3 compared to the disulfide bond constrained natural loop contributes to accommodation of target-protein interaction and allows for the successful isolation of a large set of different vNARs, albeit with low affinity to their target.

Moreover, we designed the library to be of type IV to obviate protein misfolding due to a potential presence of a significant proportion of disulfide patterns which perturb the structural integrity of the vNAR domain in a Cys-biased library. However, such semi-synthetic Cys-biased CDR3 randomized Type II libraries were constructed by Nuttall and colleagues who isolated vNAR molecules specifically targeting Gingipain K protease from *Porphyromonas ginglivalis.*[106] In contrast to our work, this group used phage-display as platform technology and nowadays, the generation of very large antibody phage-display libraries with diversities of more than 10^{11} unique clones is amenable.[195,196] This significantly exceeds diversities of yeast surface display libraries.[149,181] Thus, it can be hypothesized that it is feasible to select binders from such huge phage-display libraries that display high affinity and stability, in spite of the presence of potentially misfolded, non-functional molecules.

Nonetheless, it would be interesting to analyze the screening outcome from a vNAR library of similar diversity that contains Cys in the randomized CDR3 loop and a potentially disulfide bond-forming Cys in CDR1 with respect to structural integrity, clonal diversity of isolated binders and their target affinities.

Such Cys-biased, CDR3-diversified libraries typically give rise to binders with K_D values in the double-digit to triple-digit nanomolar range.[113] Likewise, it was shown by Goldman and co-workers, that high-affinity binders can be isolated from a semi-synthetic CDR3 mutagenized vNAR library against different antigens [114] that were lacking the inter-loop disulfide bond. In conclusion, the inter-loop disulfide bond between CDR1 and CDR3, if present, likely contributes to binding affinity and structural stability, but is not absolutely required to obtain high-affinity vNARs.

Stepwise affinity optimization by first randomizing CDR3 and accumulating a set of low affinity binders followed by diversification of CDR1 in this population of enriched variants proved a useful strategy to obtain vNARs with affinities in the low nanomolar range. As described by Wittrup and colleagues, library establishment in *S. cerevisiae* and screenable throughput limit the number of clones that can be analyzed in a combinatorial library.[156,197] As a consequence, we decided to perform this consecutive enrichment and affinity maturation procedure, to decline the theoretical sequence space in an ideally unbiased library for a more comprehensive library sampling. A simultaneous randomization of twelve residues within CDR3 and five residues in CDR1 would comprise a theoretical sequence diversity of approximately 5.5×10^{21}, whereas a consecutive diversification of twelve residues followed by randomizing five residues of drastically reduced target-enriched sequences decreases the theoretical sequence space to about 2×10^{15} and 2.5×10^{6} (the latter projected for the affinity maturation of one binder), respectively. It needs to be mentioned that even in a library of 2×10^{8} unique clones, theoretical sequence space is greatly undersampled. In contrast to the simultaneous randomization of the CDRs of the vNAR domain, this stepwise selection and

optimization technique leastwise ensures appropriate sampling of the respective sublibrary for affinity maturation that enables the analysis of every single variant in terms of enhancing affinity.

Interestingly, this novel *in vitro* affinity maturation method also resembles the natural strategy of the immune system of nurse sharks [109] to select clones from a primary IgNAR repertoire that is nearly entirely CDR3-based, followed by affinity maturation of complementarity determining regions and hypervariable loops after antigen exposure, with a significant bias towards replacement substitutions in CDR1.[98] The stepwise screening strategy is similar to the isolation of camelized human VH domains by phage display, where low-affinity binders were selected from a CDR3 randomized VH domain. Subsequent diversification of CDR1 or CDR2 residues provoked a significant increase in affinity [198,199]. While in these studies a single VH scaffold was chosen for CDR randomization, we used a repertoire of natural scaffold molecules from non-immunized animals. This may contribute to the successful isolation of a large set of stable binders, since small variations in the scaffold sequence are known to have a large impact on folding stability and protein solubility [200,201].

Although affinity matured EpCAM-binders selected with this method comprised low nanomolar affinities similar to those obtained from nurse sharks immunized with hen egg lysozyme [100], none of the selected binders showed sub-nanomolar affinities, that were recently obtained by immunization of *Squalus acanthias* with human serum albumin [101]. Nevertheless, subsequent third-round randomization of HV4 that potentially contributes to the paratope may result in variants with further enhanced affinities.

Since the development of yeast surface display, pioneered by Boder and Wittrup in 1997, it was extensively shown that this cellular display technology can be successfully utilized to engineer antibody-like molecules and also a wide range of alternative binding scaffolds [69,146,150,202-205]. As already discussed, one of the major benefits of this display method is control over the selection process, as it harbors the potential of single cell real-time and on-line analysis and subsequent characterization of individual library candidates. Another beneficial attribute is the possibility to co-select for high-level expression and stability simultaneously to the binding functionality [146,149]. Moreover, proteins are believed to be most likely well-folded since quality control machineries exist in *S. cerevisiae* for proper protein folding [206]. In addition, isolated binders can be instantaneously characterized in terms of affinity and specificity without the need for soluble expression [146,182]. Finally, sampling of libraries supposedly is more comprehensive with yeast display than with phage display [157]. This might be particularly important for the selection of rare binders from large library e.g. naïve or artificial repertoires. Yet, it needs to be mentioned, that this drawback of library sampling is at least partially compensated for phage display libraries due to the feasibility of constructing libraries with larger clone numbers.[195,196] Ultimately, this study adds vNAR antibody domains to the growing list of scaffold molecules that can be easily isolated, characterized and optimized using this powerful display technology.

A set of EpCAM-specific vNAR molecules was produced in *E. coli* as soluble protein. Though, expression yields were only moderate to low, ranging from 0.3 mg/L to 1.1 mg/L. In contrast to these findings, Nuttall and co-workers reported about vNAR yields ranging from 3 mg/L to 5 mg/L for expression into the *E. coli* periplasm.[119,125] Unlike Nuttall *et al.*, we expressed the vNAR domain as fusion protein. Consequently, two purification procedures as well as one additional manipulation step (TEV-cleavage) were performed to isolate the respective soluble domain. It can be suggested that the low expression yields are at least partially based on these multifactorial purification and manipulation procedures. It is the author's opinion, that this inherent limitation of expression yields may be optimized by obviating the production of selected vNAR molecules as fusion protein in *E. coli*. In this respect, soluble expression of the solitary vNAR domain would substantially decrease manipulation steps needed for purification and it can be concluded that this would significantly enhance expression yields.

To enhance vNAR expression, affinity-optimized α-EpCAM-vNAR 5005 was genetically fused to the *N*-terminus of a human Fc domain. The rationale was that fusions to the human Fc-part, besides enhancing vNAR expression yields, also lead to the formation of dimers. This attribute might be advantageous in terms of enhancing functional affinity i.e. avidity. Moreover, for cell-based staining assays, binding of Fc-fusion proteins to cells e.g. tumor cells can easily be detected using fluorescence-labeled secondary antibodies targeting the human Fc-part. Expression yields for α-EpCAM-vNAR 5005 Fc-fusion protein extensively exceeded those for production in *E. coli*. Thus far, the expression level of approximately 192 mg/L is to the best of our knowledge the highest reported expression yield for vNAR fusions. Typically vNARs are expressed as Fc fusion protein at levels varying from 10-40 mg/L.[103]

The superior thermal stability compared to ScFv- and mAb-formats is a hallmark of vNAR domains.[104,121] Taking a closer look in order to verify whether this is also in accordance with semi-synthetic type IV domains originating from the bamboo sharks' antibody repertoire, thermal shift assays were executed to determine the melting temperature. All three analyzed vNAR fragments exhibited high thermal stability of approximately 75 °C, 71.5 °C and 65 °C, respectively. Interestingly, affinity-optimized EpCAM-specific vNAR possessed the lowest Tm, which constitutes a significant loss of stability compared to its parental molecule H3 (Tm of ~ 75 °C). Thus, in this particular case enhancement of affinity is at the cost of thermal stability. As repeatedly mentioned, the primary repertoire of sharks is nearly exclusively CDR3-based. In conclusion, CDR1 is conserved in the IgNAR gene cluster.[24,99] Consequently, it is tempting to speculate that during evolution a very thermo-stable framework evolved and that it can be expected that mutations in CDR1, whilst increasing the affinity for antigen binding, might have a negative impact in terms of stability. Nevertheless, we could prove that even semi-synthetic type IV vNAR domains, which lack any non-canonical loop stabilizing disulfide bonds, show superior thermal stability compared to mammalian

antibody domains. Related to the latter, scFv fragments for instance, often show Tm values in the range between 50 °C and 65 °C.[207,208] However, Tm values for scFvs may sometimes fall below 40°C requiring optimization of thermal stability.[209]

As extensively reviewed, equilibrium dissociation constants measured through titration on the yeast surface are consistent with affinities determined using purified and soluble protein.[146,210,211] To examine, whether this is also consistent with vNAR domains isolated and characterized by yeast surface display, we determined affinities of EpCAM-specific vNARs from initial screening experiments and after affinity maturation *via* biolayer interferometry. Essentially, affinities for vNARs as measured on the yeast surface correlated well with dissociation constants, assessed with soluble, solitary vNAR domains. For all of the analyzed proteins the deviance between both analysis methods was less than the factor of three. This concordance is a valuable attribute for the generation of hit candidates from shark vNAR libraries using yeast surface display as platform technology, since there is no need for soluble expression and purification of individual clones in order to rank affinities of clones isolated against a given target. As also demonstrated, formatting of α-EpCAM-vNAR_5005 into the Fc-format did not impair affinities against EpCAM. Whereas we only detected a slight increase in on-rate kinetics, off-rate kinetics were significantly improved, resulting in a moderate gain of affinity, indicating a benefit due to functional affinity effects. This data suggests that vNAR domains can be engineered into the Fc-framework, without any loss of function.

Of additional concern was selectivity of isolated vNARs. For this, specificity assays were employed on the yeast surface with different unrelated target proteins for a subset of EpCAM-specific vNARs. For all tested unrelated proteins, no significant off-target binding was observed. Furthermore, for affinity-optimized α-EpCAM-vNAR_5005 we could show that there also was no unspecific binding against CD3ε and unglycosylated human Fcγ protein detectable. Contradictory to this, we observed extensive off-target binding of this vNAR variant against EpCAM-negative cell line CHO in cellular immunofluorescence assays. Conspicuously, intensities of binding for this EpCAM-negative cell line exceeded intensities for EpCAM-positive cell lines, MCF-7, EBC-1 and T47D, respectively, to a large extent. Hence, this data suggests rather specific off-target binding to a molecule of unknown nature displayed on the membrane of CHO, than unspecific off-target binding. Off-target binding is a major concern for *in vitro* and *de novo* generation of antibody-molecules against a given target, owing to the lack of any immune surveillance i.e. *in vivo* selection and regulatory mechanisms. In contrast to this, host-generated antibodies i.e. respective antibody-producing B-cells, cross-reacting to self-antigens or normal tissue are usually eliminated to prevent autoimmune disease.[212] Off-target binding can have a negative impact related to pharmacokinetics, tissue distribution, efficacy and toxicity. For instance, for adalimumab, which was generated *in vitro* from a cloned human antibody phage display library, multiple off-target interactions were found.[212,213] Importantly, it is hypothesized that serious adverse events having an autoimmune etiology can be explained by off-target activities of biologic

drugs owing to a disturbed immune balance.[213] A very elegant way to reduce off-target activities was described by Kiener and co-workers, who observed unspecific broad tissue binding during the humanization and affinity maturation process of motavizumab.[214] They identified three amino acid residues, introduced during affinity maturation that mainly mediated non-specific tissue binding. The reversion of those three residues back to the original sequence led to a reduction in off-target binding and improved overall efficacy *in vivo*. With this in mind, it can be concluded that there are important challenges that emerge beyond basic antibody-antigen binding and it seems that some of them may especially affect *in vitro* generated antibodies that suffer from the lack of regulatory immune surveillance.

Nevertheless, owing to specificity of affinity optimized α-EpCAM-vNAR_5005 with selected proteins *in vitro*, this antibody domain proved to be appropriate for further continuing proof-of-concept studies such as the generation of bi-specific vNAR molecules, where a single vNAR antibody domain simultaneously targets two different antigens, as described in the scope of this thesis.

6.2 Generation of a new antigen-binding site into the vNAR scaffold through diversification of hypervariable loop 2

Within this work, we generated a new antigen-binding site into the vNAR scaffold, solely facilitating antigen binding. Diversification of nine residues within the surface-exposed loop corresponding to hypervariable loop 2 combined with library screening using yeast surface display proved to be a valid strategy for the identification of bi-specific IgNAR V domains targeting two distinct antigens. Hence, we demonstrated that HV2 of the vNAR domain can be functionalized to form an autonomous antigen-binding site, independent from the conventional paratope, consisting of CDR3 and CDR1 (and potentially HV4).

For this, we chose to use well-characterized EpCAM-binding vNAR 5005 as scaffold for randomization. As aforementioned, this antibody domain binds to EpCAM with high affinity. Additionally, selectivity assays against unrelated target proteins demonstrated no unspecific binding against unrelated target proteins and more importantly, also no off-target binding was observed against CD3ε and non-glycosylated human Fcγ. We successfully showed, that through diversification of HV2, binders can be selected against two proof-of-concept antigens. Isolated and characterized clones comprised moderate affinities for their target, while retaining high affinity against EpCAM.

The surface-exposed loop, comprising hypervariable loop 2 is considerably longer (Lys43 – Arg56) than the nine residues, considered for library design (Ala45 – Thr 53). However, sequence comparison between different vNAR types across different species revealed a conserved Gly-Arg-Tyr motif at the end of this surface-exposed loop (Gly55 – Tyr57), as depicted in **Appendix B**. Consequently, these

residues were excluded for library design. Interestingly, CD3ε and EpCAM binding molecule B1 contained an Arg56Gln mutation. Peculiarly, this residue was not considered for library design. This molecule displayed moderate affinities against CD3ε, yet, affinities against EpCAM were substantially diminished compared to parental molecule 5005, giving clear evidence that the conserved Gly-Arg-Tyr motif is crucial for the structural integrity of the vNAR scaffold. In this respect, we were able to show that transformation into the original Gly-Arg-Tyr motif entirely restored affinities towards EpCAM without any loss of function against CD3ε.

Characterization of the established library by FACS analysis for staining with EpCAM showed that the vast majority of displayed clones (~ 89.5 %) retained high affinity for the original antigen, indicating that a large portion of library candidates tolerated well the modification of this structural loop. Related to this, isolated bi-specific vNAR clones did not display any significant loss of affinity against EpCAM, demonstrating the versatility of this new strategy.

It is generally known, that the immunoglobulin family evolved a paramount tolerability in loop length as well as sequence variation, which is most evident for CDR regions of variable domains. However, this feature is a general hallmark of the immunoglobulin domains.[69,215] In this respect, surface-exposed loops at the *N*-terminal tip of CH2 of human IgG have been engineered for antigen binding.[216] In another very elegant approach, established by Rüker and co-workers, a new antigen-binding site was introduced into *C*-terminal loop regions of the CH3 domain of human IgG.[69] Most importantly, those engineered Fc fragments, named Fcab, retained the ability to elicit effector functions, clearly indicating the structural integrity of the molecule. This study adds the vNAR domain to the growing class of immunoglobulin scaffold proteins that can be manipulated and tailor-made in terms of establishing a new antigen-binding site.

Although this new methodology allows for the identification and isolation of binders comprising a new antigen-binding site, hence, for bi-specific molecules, affinities of the new HV2-facilitated paratope against their target proteins were only moderate. Notwithstanding, there might be several methodologies applicable to optimize affinities. One could think of randomizing residues Lys43 and Gly44 of HV2 to expand the created antigen-binding loop which may contribute to an enhanced affinity i.e. to affinity maturation. Another possibility might be to identify residues of HV2 not involved in antigen binding *via* alanine-scanning. After identification, those residues could be diversified in a second generation randomization step and a sublibrary could be established. Screening with significantly decreased target concentration, as performed in the aforementioned affinity maturation process of the conventional antigen-binding site, i.e. randomization of CDR1, might pave the way for the identification and isolation of bi-specific vNAR molecules with significantly enhanced affinities. Finally, randomization of the adjacent loop of HV2 (Asp74 – Ser80) would substantially elongate the surface of a potentially antigen-binding site, drastically increasing the chance to isolate high-affinity binders, which would be progenies of HV2-randomized antigen binding, bi-specific

clones. Accordingly, this antigen-binding site would comprise two surface-exposed loops, akin to the conventional paratope of EpCAM-binding high-affinity vNAR 5005, consisting of CDR3 and CDR1. Such antigen-binding sites were constructed into domain CH3 of the human Fc-fragment of IgG, as described by Rüker and colleagues.[69] The group randomized the AB and EF loop at the *C*-terminal tip of CH3 and screened for antigen binding Fc fragments using yeast surface display. With this strategy, comprising a two-looped paratope, they were able to isolate Fc fragments targeting HER2/neu with affinities of 69 nM. Interestingly, also insertions into the surface-exposed loops considered for library design did not interfere with the overall structural and functional characteristics of the Fc-part. As a consequence, one could also take insertions of randomized residues into HV2 into account to elongate the respective loop, for the enhancement of affinities.

Essentially, this study demonstrates that the vNAR scaffold can be engineered in a way that HV2 functions as an independent paratope, solely facilitating antigen binding against a target protein. Importantly, the establishment of a new paratope does not impair the conventional antigen-binding site, composed of CDR3 and CDR1 in its affinity, resulting in a bi-specfifc molecule.

In further studies, taking a closer look to bi-specificity might be worth to consider. In this respect, vNAR molecules need to be expressed as soluble proteins and characterized more meticulously *via* biolayer interferometry to investigate whether the vNAR molecule is capable of targeting both antigens simultaneously. Although HV2 of the IgNAR V domain is situated in relative distance to the conventional paratope, with the data presented herein, it cannot be excluded that both antigen-binding sites compete for antigen binding to their respective target proteins, which would prevent simultaneous targeting.

7. REFERENCES

1. Medzhitov R, Janeway C, Jr.: **Innate immune recognition: Mechanisms and pathways.** *Immunol Rev* (2000) **173**(89-97.

2. Jefferis R: **Isotype and glycoform selection for antibody therapeutics.** *Archives of biochemistry and biophysics* (2012) **526**(2):159-166.

3. Schenten D, Medzhitov R: **The control of adaptive immune responses by the innate immune system.** *Advances in immunology* (2011) **109**(87-124.

4. Iwasaki A, Medzhitov R: **Regulation of adaptive immunity by the innate immune system.** *Science* (2010) **327**(5963):291-295.

5. Janeway CA, Jr., Medzhitov R: **Innate immune recognition.** *Annual review of immunology* (2002) **20**(197-216.

6. Buchmann K: **Evolution of innate immunity: Clues from invertebrates via fish to mammals.** *Frontiers in immunology* (2014) **5**(459.

7. Medzhitov R, Janeway CA, Jr.: **Innate immunity: Impact on the adaptive immune response.** *Curr Opin Immunol* (1997) **9**(1):4-9.

8. Akira S, Uematsu S, Takeuchi O: **Pathogen recognition and innate immunity.** *Cell* (2006) **124**(4):783-801.

9. Flajnik MF, Du Pasquier L: **Evolution of innate and adaptive immunity: Can we draw a line?** *Trends Immunol* (2004) **25**(12):640-644.

10. Cooper MD, Alder MN: **The evolution of adaptive immune systems.** *Cell* (2006) **124**(4):815-822.

11. Ghosh J, Lun CM, Majeske AJ, Sacchi S, Schrankel CS, Smith LC: **Invertebrate immune diversity.** *Developmental and comparative immunology* (2011) **35**(9):959-974.

12. Watson FL, Puttmann-Holgado R, Thomas F, Lamar DL, Hughes M, Kondo M, Rebel VI, Schmucker D: **Extensive diversity of Ig-superfamily proteins in the immune system of insects.** *Science* (2005) **309**(5742):1874-1878.

13. Zhang SM, Adema CM, Kepler TB, Loker ES: **Diversification of Ig superfamily genes in an invertebrate.** *Science* (2004) **305**(5681):251-254.

14. Kasahara M, Sutoh Y: **Two forms of adaptive immunity in vertebrates: Similarities and differences.** *Advances in immunology* (2014) **122**(59-90.

15. Pancer Z, Amemiya CT, Ehrhardt GR, Ceitlin J, Gartland GL, Cooper MD: **Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey.** *Nature* (2004) **430**(6996):174-180.

16. Pancer Z, Saha NR, Kasamatsu J, Suzuki T, Amemiya CT, Kasahara M, Cooper MD: **Variable lymphocyte receptors in hagfish.** *Proceedings of the National Academy of Sciences of the United States of America* (2005) **102**(26):9224-9229.

17. Boehm T, McCurley N, Sutoh Y, Schorpp M, Kasahara M, Cooper MD: **VLR-based adaptive immunity.** *Annual review of immunology* (2012) **30**(203-220.

18. Wesolowski J, Alzogaray V, Reyelt J, Unger M, Juarez K, Urrutia M, Cauerhff A, Danquah W, Rissiek B, Scheuplein F, Schwarz N *et al*: **Single domain antibodies: Promising experimental and therapeutic tools in infection and immunity.** *Med Microbiol Immunol* (2009) **198**(3):157-174.

19. Wang F, Ekiert DC, Ahmad I, Yu W, Zhang Y, Bazirgan O, Torkamani A, Raudsepp T, Mwangi W, Criscitiello MF, Wilson IA *et al*: **Reshaping antibody diversity.** *Cell* (2013) **153**(6):1379-1393.

20. Hamers-Casterman C, Atarhouch T, Muyldermans S, Robinson G, Hamers C, Songa EB, Bendahman N, Hamers R: **Naturally occurring antibodies devoid of light chains.** *Nature* (1993) **363**(6428):446-448.

21. Greenberg AS, Avila D, Hughes M, Hughes A, McKinney EC, Flajnik MF: **A new antigen receptor gene family that undergoes rearrangement and extensive somatic diversification in sharks.** *Nature* (1995) **374**(6518):168-173.

22. Flajnik MF, Deschacht N, Muyldermans S: **A case of convergence: Why did a simple alternative to canonical antibodies arise in sharks and camels?** *PLoS biology* (2011) **9**(8):e1001120.

23. Desmyter A, Transue TR, Ghahroudi MA, Thi MH, Poortmans F, Hamers R, Muyldermans S, Wyns L: **Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme.** *Nature structural biology* (1996) **3**(9):803-811.

24. Stanfield RL, Dooley H, Flajnik MF, Wilson IA: **Crystal structure of a shark single-domain antibody V region in complex with lysozyme.** *Science* (2004) **305**(5691):1770-1773.

25. Streltsov VA, Carmichael JA, Nuttall SD: **Structure of a shark IgNAR antibody variable domain and modeling of an early-developmental isotype.** *Protein science : a publication of the Protein Society* (2005) **14**(11):2901-2909.

26. Holz JB: **The titan trial--assessing the efficacy and safety of an anti-von willebrand factor nanobody in patients with acquired thrombotic thrombocytopenic purpura.** *Transfusion and apheresis science : official journal of the World Apheresis Association : official journal of the European Society for Haemapheresis* (2012) **46**(3):343-346.

27. Adderson EE, Johnston JM, Shackelford PG, Carroll WL: **Development of the human antibody repertoire.** *Pediatric research* (1992) **32**(3):257-263.

28. Di Noia JM, Neuberger MS: **Molecular mechanisms of antibody somatic hypermutation.** *Annual review of biochemistry* (2007) **76**(1-22.

29. Longerich S, Storb U: **The contested role of uracil DNA glycosylase in immunoglobulin gene diversification.** *Trends in genetics : TIG* (2005) **21**(5):253-256.

30. Stavnezer J, Guikema JE, Schrader CE: **Mechanism and regulation of class switch recombination.** *Annual review of immunology* (2008) **26**(261-292.

31. Schroeder HW, Jr., Cavacini L: **Structure and function of immunoglobulins.** *The Journal of allergy and clinical immunology* (2010) **125**(2 Suppl 2):S41-52.

32. Nezlin R, Ghetie V: **Interactions of immunoglobulins outside the antigen-combining site.** *Advances in immunology* (2004) **82**(155-215.

33. An Z: **Monoclonal antibodies - a proven and rapidly expanding therapeutic modality for human diseases.** *Protein & cell* (2010) **1**(4):319-330.

34. Buss NA, Henderson SJ, McFarlane M, Shenton JM, de Haan L: **Monoclonal antibody therapeutics: History and future.** *Current opinion in pharmacology* (2012) **12**(5):615-622.

35. Nelson AL, Dhimolea E, Reichert JM: **Development trends for human monoclonal antibody therapeutics.** *Nat Rev Drug Discov* (2010) **9**(10):767-774.

36. Yang J, Sundrud MS, Skepner J, Yamagata T: **Targeting Th17 cells in autoimmune diseases.** *Trends Pharmacol Sci* (2014) **35**(10):493-500.

37. Lee JK, Joo KM, Lee J, Yoon Y, Nam DH: **Targeting the epithelial to mesenchymal transition in glioblastoma: The emerging role of met signaling.** *OncoTargets and therapy* (2014) **7**(1933-1944.

38. Harris LJ, Larson SB, Hasel KW, McPherson A: **Refined structure of an intact IgG2a monoclonal antibody.** *Biochemistry* (1997) **36**(7):1581-1597.

39. Zielonka S, Empting M, Grzeschik J, Konning D, Barelle CJ, Kolmar H: **Structural insights and biomedical potential of IgNAR scaffolds from sharks.** *mAbs* (2015) **7**(1):15-25.

40. Takai T: **Roles of Fc receptors in autoimmunity.** *Nature reviews Immunology* (2002) **2**(8):580-592.

41. Ravetch JV, Bolland S: **IgG Fc receptors.** *Annual review of immunology* (2001) **19**(275-290.

42. Tamm A, Schmidt RE: **IgG binding sites on human Fc gamma receptors.** *International reviews of immunology* (1997) **16**(1-2):57-85.

43. Ghetie V, Ward ES: **Multiple roles for the major histocompatibility complex class I- related receptor FcRn.** *Annual review of immunology* (2000) **18**(739-766.

44. Idusogie EE, Presta LG, Gazzano-Santoro H, Totpal K, Wong PY, Ultsch M, Meng YG, Mulkerrin MG: **Mapping of the C1q binding site on rituxan, a chimeric antibody with a human IgG1 Fc.** *J Immunol* (2000) **164**(8):4178-4184.

45. Aggarwal RS: **What's fueling the biotech engine-2012 to 2013.** *Nat Biotechnol* (2014) **32**(1):32-39.

46. Aggarwal SR: **What's fueling the biotech engine-2011 to 2012.** *Nature biotechnology* (2012) **30**(12):1191-1197.

47. Aggarwal S: **What's fueling the biotech engine--2010 to 2011.** *Nature biotechnology* (2011) **29**(12):1083-1089.

48. Elvin JG, Couston RG, van der Walle CF: **Therapeutic antibodies: Market considerations, disease targets and bioprocessing.** *International journal of pharmaceutics* (2013) **440**(1):83-98.

49. Reichert JM: **Marketed therapeutic antibodies compendium.** *mAbs* (2012) **4**(3):413-415.

50. Reichert JM: **Antibodies to watch in 2014: Mid-year update.** *mAbs* (2014) **6**(4):799-802.

51. Barelle C, Gill DS, Charlton K: **Shark novel antigen receptors--the next generation of biologic therapeutics?** *Adv Exp Med Biol* (2009) **655**(49-62.

52. Nuttall SD, Walsh RB: **Display scaffolds: Protein engineering for novel therapeutics.** *Current opinion in pharmacology* (2008) **8**(5):609-615.

53. Mordenti J, Cuthbertson RA, Ferrara N, Thomsen K, Berleau L, Licko V, Allen PC, Valverde CR, Meng YG, Fei DT, Fourre KM *et al*: **Comparisons of the intraocular tissue distribution, pharmacokinetics, and safety of 125i-labeled full-length and Fab antibodies in rhesus monkeys following intravitreal administration.** *Toxicologic pathology* (1999) **27**(5):536-544.

54. Vaneycken I, D'Huyvetter M, Hernot S, De Vos J, Xavier C, Devoogdt N, Caveliers V, Lahoutte T: **Immuno-imaging using nanobodies.** *Curr Opin Biotechnol* (2011) **22**(6):877-881.

55. Banta S, Dooley K, Shur O: **Replacing antibodies: Engineering new binding proteins.** *Annual review of biomedical engineering* (2013) **15**(93-113.

56. Huang L, Gainkam LO, Caveliers V, Vanhove C, Keyaerts M, De Baetselier P, Bossuyt A, Revets H, Lahoutte T: **Spect imaging with 99mtc-labeled EGFR-specific nanobody for in vivo monitoring of EGFR expression.** *Molecular imaging and biology : MIB : the official publication of the Academy of Molecular Imaging* (2008) **10**(3):167-175.

57. Kenanova V, Wu AM: **Tailoring antibodies for radionuclide delivery.** *Expert opinion on drug delivery* (2006) **3**(1):53-70.

58. Van de Wiele C, Revets H, Mertens N: **Radioimmunoimaging. Advances and prospects.** *The quarterly journal of nuclear medicine and molecular imaging : official publication of the Italian Association of Nuclear Medicine (AIMN) [and] the International Association of Radiopharmacology (IAR), [and] Section of the So* (2004) **48**(4):317-325.

59. Kontermann RE: **Dual targeting strategies with bispecific antibodies.** *mAbs* (2012) **4**(2):182-197.

60. Morgillo F, Lee HY: **Resistance to epidermal growth factor receptor-targeted therapy.** *Drug resistance updates : reviews and commentaries in antimicrobial and anticancer chemotherapy* (2005) **8**(5):298-310.

61. Elia G, Fugmann T, Neri D: **From target discovery to clinical trials with armed antibody products.** *Journal of proteomics* (2014) **107c**(50-55.

62. Enever C, Batuwangala T, Plummer C, Sepp A: **Next generation immunotherapeutics--honing the magic bullet.** *Current opinion in biotechnology* (2009) **20**(4):405-411.

63. Gebauer M, Skerra A: **Engineered protein scaffolds as next-generation antibody therapeutics.** *Current opinion in chemical biology* (2009) **13**(3):245-255.

64. Holt LJ, Herring C, Jespers LS, Woolven BP, Tomlinson IM: **Domain antibodies: Proteins for therapy.** *Trends in biotechnology* (2003) **21**(11):484-490.

65. Lambert JM: **Drug-conjugated antibodies for the treatment of cancer.** *British journal of clinical pharmacology* (2013) **76**(2):248-262.

66. Lofblom J, Frejd FY, Stahl S: **Non-immunoglobulin based protein scaffolds.** *Current opinion in biotechnology* (2011) **22**(6):843-848.

67. Muller D, Kontermann RE: **Bispecific antibodies for cancer immunotherapy: Current perspectives.** *BioDrugs : clinical immunotherapeutics, biopharmaceuticals and gene therapy* (2010) **24**(2):89-98.

68. Nelson AL: **Antibody fragments: Hope and hype.** *MAbs* (2010) **2**(1):77-83.

69. Wozniak-Knopp G, Bartl S, Bauer A, Mostageer M, Woisetschlager M, Antes B, Ettl K, Kainer M, Weberhofer G, Wiederkum S, Himmler G *et al*: **Introducing antigen-binding sites in structural loops of immunoglobulin constant domains: Fc fragments with engineered HER2/neu-binding sites and antibody properties.** *Protein engineering, design & selection : PEDS* (2010) **23**(4):289-297.

70. Kolmar H: **Natural and engineered cystine knot miniproteins for diagnostic and therapeutic applications.** *Curr Pharm Des* (2011) **17**(38):4329-4336.

71. Byrne H, Conroy PJ, Whisstock JC, O'Kennedy RJ: **A tale of two specificities: Bispecific antibodies for therapeutic and diagnostic applications.** *Trends in biotechnology* (2013) **31**(11):621-632.

72. Demarest SJ, Hariharan K, Dong J: **Emerging antibody combinations in oncology.** *mAbs* (2011) **3**(4):338-351.

73. Wheeler DL, Huang S, Kruser TJ, Nechrebecki MM, Armstrong EA, Benavente S, Gondi V, Hsu KT, Harari PM: **Mechanisms of acquired resistance to Cetuximab: Role of her (ErbB) family members.** *Oncogene* (2008) **27**(28):3944-3956.

74. Riedemann J, Takiguchi M, Sohail M, Macaulay VM: **The EGF receptor interacts with the type 1 IgF receptor and regulates its stability.** *Biochem Biophys Res Commun* (2007) **355**(3):707-714.

75. Hendrickson AW, Haluska P: **Resistance pathways relevant to insulin-like growth factor-1 receptor-targeted therapy.** *Current opinion in investigational drugs (London, England : 2000)* (2009) **10**(10):1032-1040.

76. Zhu Z, Yan L: **Next generation of antibody therapy for cancer.** *Chinese journal of cancer* (2011) **30**(5):293-302.

77. Beck A, Wurch T, Bailly C, Corvaia N: **Strategies and challenges for the next generation of therapeutic antibodies.** *Nature reviews Immunology* (2010) **10**(5):345-352.

78. Klein C, Sustmann C, Thomas M, Stubenrauch K, Croasdale R, Schanzer J, Brinkmann U, Kettenberger H, Regula JT, Schaefer W: **Progress in overcoming the chain association issue in bispecific heterodimeric IgG antibodies.** *mAbs* (2012) **4**(6):653-663.

79. Chames P, Baty D: **Bispecific antibodies for cancer therapy: The light at the end of the tunnel?** *mAbs* (2009) **1**(6):539-547.

80. May C, Sapra P, Gerber HP: **Advances in bispecific biotherapeutics for the treatment of cancer.** *Biochem Pharmacol* (2012) **84**(9):1105-1112.

81. Linke R, Klein A, Seimetz D: **Catumaxomab: Clinical development and future directions.** *mAbs* (2010) **2**(2):129-136.

82. Lameris R, de Bruin RC, Schneiders FL, van Bergen En Henegouwen PM, Verheul HM, de Gruijl TD, van der Vliet HJ: **Bispecific antibody platforms for cancer immunotherapy.** *Crit Rev Oncol Hematol* (2014).

83. Ridgway JB, Presta LG, Carter P: **'Knobs-into-holes' engineering of antibody CH3 domains for heavy chain heterodimerization.** *Protein engineering* (1996) **9**(7):617-621.

84. Muda M, Gross AW, Dawson JP, He C, Kurosawa E, Schweickhardt R, Dugas M, Soloviev M, Bernhardt A, Fischer D, Wesolowski JS *et al*: **Therapeutic assessment of SEED: A new engineered antibody platform designed to generate mono- and bispecific antibodies.** *Protein engineering, design & selection : PEDS* (2011) **24**(5):447-454.

85. Wranik BJ, Christensen EL, Schaefer G, Jackman JK, Vendel AC, Eaton D: **Luz-y, a novel platform for the mammalian cell production of full-length IgG-bispecific antibodies*.** *The Journal of biological chemistry* (2012) **287**(52):43331-43339.

86. Schaefer W, Regula JT, Bahner M, Schanzer J, Croasdale R, Durr H, Gassner C, Georges G, Kettenberger H, Imhof-Jung S, Schwaiger M *et al*: **Immunoglobulin domain crossover as a generic approach for the production of bispecific IgG antibodies.** *Proceedings of the National Academy of Sciences of the United States of America* (2011) **108**(27):11187-11192.

87. Rumfelt LL, Lohr RL, Dooley H, Flajnik MF: **Diversity and repertoire of IgW and IgM VH families in the newborn nurse shark.** *BMC immunology* (2004) **5**(8.

88. Rumfelt LL, Diaz M, Lohr RL, Mochon E, Flajnik MF: **Unprecedented multiplicity of Ig transmembrane and secretory mRNA forms in the cartilaginous fish.** *Journal of immunology (Baltimore, Md : 1950)* (2004) **173**(2):1129-1139.

89. Greenberg AS, Hughes AL, Guo J, Avila D, McKinney EC, Flajnik MF: **A novel "chimeric" antibody class in cartilaginous fish: IgM may not be the primordial immunoglobulin.** *European journal of immunology* (1996) **26**(5):1123-1129.

90. Dooley H, Flajnik MF: **Antibody repertoire development in cartilaginous fish.** *Developmental and comparative immunology* (2006) **30**(1-2):43-56.

91. Feige MJ, Grawert MA, Marcinowski M, Hennig J, Behnke J, Auslander D, Herold EM, Peschek J, Castro CD, Flajnik M, Hendershot LM *et al*: **The structural analysis of shark IgNAR antibodies reveals evolutionary principles of immunoglobulins.** *Proceedings of the National Academy of Sciences of the United States of America* (2014) **111**(22):8155-8160.

92. Roux KH, Greenberg AS, Greene L, Strelets L, Avila D, McKinney EC, Flajnik MF: **Structural analysis of the nurse shark (new) antigen receptor (NAR): Molecular convergence of NAR and unusual mammalian immunoglobulins.** *Proceedings of the National Academy of Sciences of the United States of America* (1998) **95**(20):11804-11809.

93. Feige MJ, Groscurth S, Marcinowski M, Shimizu Y, Kessler H, Hendershot LM, Buchner J: **An unfolded CH1 domain controls the assembly and secretion of IgG antibodies.** *Molecular cell* (2009) **34**(5):569-579.

94. Hsu E, Pulham N, Rumfelt LL, Flajnik MF: **The plasticity of immunoglobulin gene systems in evolution.** *Immunological reviews* (2006) **210**(8-26.

95. Berstein RM, Schluter SF, Shen S, Marchalonis JJ: **A new high molecular weight immunoglobulin class from the carcharhine shark: Implications for the properties of the primordial immunoglobulin.** *Proceedings of the National Academy of Sciences of the United States of America* (1996) **93**(8):3289-3293.

96. Criscitiello MF, Saltis M, Flajnik MF: **An evolutionarily mobile antigen receptor variable region gene: Doubly rearranging NAR-TCR genes in sharks.** *Proceedings of the National Academy of Sciences of the United States of America* (2006) **103**(13):5036-5041.

97. Streltsov VA, Varghese JN, Carmichael JA, Irving RA, Hudson PJ, Nuttall SD: **Structural evidence for evolution of shark Ig new antigen receptor variable domain antibodies from a cell-surface receptor.** *Proceedings of the National Academy of Sciences of the United States of America* (2004) **101**(34):12444-12449.

98. Dooley H, Stanfield RL, Brady RA, Flajnik MF: **First molecular and biochemical analysis of in vivo affinity maturation in an ectothermic vertebrate.** *Proceedings of the National Academy of Sciences of the United States of America* (2006) **103**(6):1846-1851.

99. Stanfield RL, Dooley H, Verdino P, Flajnik MF, Wilson IA: **Maturation of shark single-domain (IgNAR) antibodies: Evidence for induced-fit binding.** *Journal of molecular biology* (2007) **367**(2):358-372.

100. Dooley H, Flajnik MF, Porter AJ: **Selection and characterization of naturally occurring single-domain (IgNAR) antibody fragments from immunized sharks by phage display.** *Molecular immunology* (2003) **40**(1):25-33.

101. Muller MR, Saunders K, Grace C, Jin M, Piche-Nicholas N, Steven J, O'Dwyer R, Wu L, Khetemenee L, Vugmeyster Y, Hickling TP *et al*: **Improving the pharmacokinetic properties of biologics by fusion to an anti-HSA shark vNAR domain.** *mAbs* (2012) **4**(6):673-685.

102. Diaz M, Stanfield RL, Greenberg AS, Flajnik MF: **Structural analysis, selection, and ontogeny of the shark new antigen receptor (IgNAR): Identification of a new locus preferentially expressed in early development.** *Immunogenetics* (2002) **54**(7):501-512.

103. Kovalenko OV, Olland A, Piche-Nicholas N, Godbole A, King D, Svenson K, Calabro V, Muller MR, Barelle CJ, Somers W, Gill DS *et al*: **Atypical antigen recognition mode of a shark immunoglobulin new antigen receptor (IgNAR) variable domain characterized by humanization and structural analysis.** *The Journal of biological chemistry* (2013) **288**(24):17408-17419.

104. Liu JL, Anderson GP, Delehanty JB, Baumann R, Hayhurst A, Goldman ER: **Selection of cholera toxin specific IgNAR single-domain antibodies from a naive shark library.** *Molecular immunology* (2007) **44**(7):1775-1783.

105. Kovaleva M, Ferguson L, Steven J, Porter A, Barelle C: **Shark variable new antigen receptor biologics - a novel technology platform for therapeutic drug development.** *Expert opinion on biological therapy* (2014) 1-13.

106. Nuttall SD, Krishnan UV, Hattarki M, De Gori R, Irving RA, Hudson PJ: **Isolation of the new antigen receptor from wobbegong sharks, and use as a scaffold for the display of protein loop libraries.** *Molecular immunology* (2001) **38**(4):313-326.

107. Krieger E, Joo K, Lee J, Lee J, Raman S, Thompson J, Tyka M, Baker D, Karplus K: **Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in casp8.** *Proteins* (2009) **77 Suppl 9**(114-122.

108. Jackson KJ, Kidd MJ, Wang Y, Collins AM: **The shape of the lymphocyte receptor repertoire: Lessons from the B cell receptor.** *Frontiers in immunology* (2013) **4**(263.

109. Diaz M, Greenberg AS, Flajnik MF: **Somatic hypermutation of the new antigen receptor gene (NAR) in the nurse shark does not generate the repertoire: Possible role in antigen-driven reactions in the absence of germinal centers.** *Proceedings of the National Academy of Sciences of the United States of America* (1998) **95**(24):14343-14348.

110. Diaz M, Velez J, Singh M, Cerny J, Flajnik MF: **Mutational pattern of the nurse shark antigen receptor gene (NAR) is similar to that of mammalian ig genes and to spontaneous mutations in evolution: The translesion synthesis model of somatic hypermutation.** *Int Immunol* (1999) **11**(5):825-833.

111. Rumfelt LL, Avila D, Diaz M, Bartl S, McKinney EC, Flajnik MF: **A shark antibody heavy chain encoded by a nonsomatically rearranged VDJ is preferentially expressed in early development and is convergent with mammalian IgG.** *Proceedings of the National Academy of Sciences of the United States of America* (2001) **98**(4):1775-1780.

112. Smith LE, Crouch K, Cao W, Muller MR, Wu L, Steven J, Lee M, Liang M, Flajnik MF, Shih HH, Barelle CJ *et al*: **Characterization of the immunoglobulin repertoire of the spiny dogfish (*Squalus acanthias*).** *Developmental and comparative immunology* (2012) **36**(4):665-679.

113. Nuttall SD, Krishnan UV, Doughty L, Nathanielsz A, Ally N, Pike RN, Hudson PJ, Kortt AA, Irving RA: **A naturally occurring NAR variable domain binds the kgp protease from *Porphyromonas gingivalis*.** *FEBS Lett* (2002) **516**(1-3):80-86.

114. Liu JL, Anderson GP, Goldman ER: **Isolation of anti-toxin single domain antibodies from a semi-synthetic spiny dogfish shark display library.** *BMC Biotechnol* (2007) **7**(78.

115. Ohtani M, Hikima J, Jung TS, Kondo H, Hirono I, Aoki T: **Construction of an artificially randomized IgNAR phage display library: Screening of variable regions that bind to hen egg white lysozyme.** *Marine biotechnology (New York, NY)* (2013) **15**(1):56-62.

116. Ohtani M, Hikima J, Jung TS, Kondo H, Hirono I, Takeyama H, Aoki T: **Variable domain antibodies specific for viral hemorrhagic septicemia virus (VHSV) selected from a randomized IgNAR phage display library.** *Fish & shellfish immunology* (2013) **34**(2):724-728.

117. Kopsidas G, Roberts AS, Coia G, Streltsov VA, Nuttall SD: **In vitro improvement of a shark IgNAR antibody by Qβ replicase mutation and ribosome display mimics in vivo affinity maturation.** *Immunol Lett* (2006) **107**(2):163-168.

118. Shao CY, Secombes CJ, Porter AJ: **Rapid isolation of IgNAR variable single-domain antibody fragments from a shark synthetic library.** *Molecular immunology* (2007) **44**(4):656-665.

119. Nuttall SD, Krishnan UV, Doughty L, Pearson K, Ryan MT, Hoogenraad NJ, Hattarki M, Carmichael JA, Irving RA, Hudson PJ: **Isolation and characterization of an IgNAR variable domain specific for the human mitochondrial translocase receptor TOM70.** *European journal of biochemistry / FEBS* (2003) **270**(17):3543-3554.

120. Camacho-Villegas T, Mata-Gonzalez T, Paniagua-Solis J, Sanchez E, Licea A: **Human tnf cytokine neutralization with a vNAR from *Heterodontus francisci* shark: A potential therapeutic use.** *mAbs* (2013) **5**(1):80-85.

121. Goodchild SA, Dooley H, Schoepp RJ, Flajnik M, Lonsdale SG: **Isolation and characterisation of Ebolavirus-specific recombinant antibody fragments from murine and shark immune libraries.** *Molecular immunology* (2011) **48**(15-16):2027-2037.

122. Dooley H, Flajnik MF: **Shark immunity bites back: Affinity maturation and memory response in the nurse shark, *Ginglymostoma cirratum*.** *European journal of immunology* (2005) **35**(3):936-945.

123. Crouch K, Smith LE, Williams R, Cao W, Lee M, Jensen A, Dooley H: **Humoral immune response of the small-spotted catshark, *Scyliorhinus canicula*.** *Fish & shellfish immunology* (2013) **34**(5):1158-1169.

124. Walsh R, Nuttall S, Revill P, Colledge D, Cabuang L, Soppe S, Dolezal O, Griffiths K, Bartholomeusz A, Locarnini S: **Targeting the Hepatitis B virus precore antigen with a novel IgNAR single variable domain intrabody.** *Virology* (2011) **411**(1):132-141.

125. Nuttall SD, Humberstone KS, Krishnan UV, Carmichael JA, Doughty L, Hattarki M, Coley AM, Casey JL, Anders RF, Foley M, Irving RA *et al*: **Selection and affinity maturation of IgNAR variable domains targeting *Plasmodium falciparum* AMA1.** *Proteins* (2004) **55**(1):187-197.

126. Simmons DP, Streltsov VA, Dolezal O, Hudson PJ, Coley AM, Foley M, Proll DF, Nuttall SD: **Shark IgNAR antibody mimotopes target a murine immunoglobulin through extended CDR3 loop structures.** *Proteins* (2008) **71**(1):119-130.

127. Henderson KA, Streltsov VA, Coley AM, Dolezal O, Hudson PJ, Batchelor AH, Gupta A, Bai T, Murphy VJ, Anders RF, Foley M *et al*: **Structure of an IgNAR-AMA1 complex: Targeting a conserved hydrophobic cleft broadens malarial strain recognition.** *Structure (London, England : 1993)* (2007) **15**(11):1452-1466.

128. Goodchild SA, Dooley H, Schoepp RJ, Flajnik M, Lonsdale SG: **Isolation and characterisation of Ebolavirus-specific recombinant antibody fragments from murine and shark immune libraries.** *Mol Immunol* (2011) **48**(15–16):2027-2037.

129. Streltsov VA, Varghese JN, Masters CL, Nuttall SD: **Crystal structure of the amyloid-beta p3 fragment provides a model for oligomer formation in alzheimer's disease.** *The Journal of neuroscience : the official journal of the Society for Neuroscience* (2011) **31**(4):1419-1426.

130. Bojalil R, Mata-Gonzalez MT, Sanchez-Munoz F, Yee Y, Argueta I, Bolanos L, Amezcua-Guerra LM, Camacho-Villegas TA, Sanchez-Castrejon E, Garcia-Ubbelohde WJ, Licea-Navarro AF *et al*: **Anti-tumor necrosis factor vNAR single domains reduce lethality and regulate underlying inflammatory response in a murine model of endotoxic shock.** *BMC Immunol* (2013) **14**(17.

131. Simmons DP, Abregu FA, Krishnan UV, Proll DF, Streltsov VA, Doughty L, Hattarki MK, Nuttall SD: **Dimerisation strategies for shark IgNAR single domain antibody fragments.** *Journal of immunological methods* (2006) **315**(1-2):171-184.

132. Muller MR, O'Dwyer R, Kovaleva M, Rudkin F, Dooley H, Barelle CJ: **Generation and isolation of target-specific single-domain antibodies from shark immune repertoires.** *Methods in molecular biology* (2012) **907**(177-194.

133. Lu ZJ, Deng SJ, Huang DG, He Y, Lei M, Zhou L, Jin P: **Frontier of therapeutic antibody discovery: The challenges and how to face them.** *World journal of biological chemistry* (2012) **3**(12):187-196.

134. Osbourn J, Groves M, Vaughan T: **From rodent reagents to human therapeutics using antibody guided selection.** *Methods (San Diego, Calif)* (2005) **36**(1):61-68.

135. De Pascalis R, Iwahashi M, Tamura M, Padlan EA, Gonzales NR, Santos AD, Giuliano M, Schuck P, Schlom J, Kashmiri SV: **Grafting of "abbreviated" complementarity-determining regions containing specificity-determining residues essential for ligand contact to engineer a less immunogenic humanized monoclonal antibody.** *Journal of immunology (Baltimore, Md : 1950)* (2002) **169**(6):3076-3084.

136. Ewert S, Honegger A, Pluckthun A: **Stability improvement of antibodies for extracellular and intracellular applications: CDR grafting to stable frameworks and structure-based framework engineering.** *Methods (San Diego, Calif)* (2004) **34**(2):184-199.

137. Roguska MA, Pedersen JT, Keddy CA, Henry AH, Searle SJ, Lambert JM, Goldmacher VS, Blattler WA, Rees AR, Guild BC: **Humanization of murine monoclonal antibodies through variable**

domain resurfacing. *Proceedings of the National Academy of Sciences of the United States of America* (1994) **91**(3):969-973.

138. Riechmann L, Clark M, Waldmann H, Winter G: **Reshaping human antibodies for therapy.** *Nature* (1988) **332**(6162):323-327.

139. Riechmann L, Muyldermans S: **Single domain antibodies: Comparison of camel VH and camelised human VH domains.** *J Immunol Methods* (1999) **231**(1-2):25-38.

140. Uth C, Zielonka S, Horner S, Rasche N, Plog A, Orelma H, Avrutina O, Zhang K, Kolmar H: **A chemoenzymatic approach to protein immobilization onto crystalline cellulose nanoscaffolds.** *Angewandte Chemie (International ed in English)* (2014).

141. Kohler G, Milstein C: **Continuous cultures of fused cells secreting antibody of predefined specificity.** *Nature* (1975) **256**(5517):495-497.

142. Sergeeva A, Kolonin MG, Molldrem JJ, Pasqualini R, Arap W: **Display technologies: Application for the discovery of drug and gene delivery agents.** *Adv Drug Deliv Rev* (2006) **58**(15):1622-1654.

143. Jakobovits A: **Production of fully human antibodies by transgenic mice.** *Curr Opin Biotechnol* (1995) **6**(5):561-566.

144. Paschke M: **Phage display systems and their applications.** *Appl Microbiol Biotechnol* (2006) **70**(1):2-11.

145. Uchiyama F, Tanaka Y, Minari Y, Tokui N: **Designing scaffolds of peptides for phage display libraries.** *J Biosci Bioeng* (2005) **99**(5):448-456.

146. Gai SA, Wittrup KD: **Yeast surface display for protein engineering and characterization.** *Current opinion in structural biology* (2007) **17**(4):467-473.

147. Wittrup KD: **Protein engineering by cell-surface display.** *Curr Opin Biotechnol* (2001) **12**(4):395-399.

148. Thie H, Meyer T, Schirrmann T, Hust M, Dubel S: **Phage display derived therapeutic antibodies.** *Curr Pharm Biotechnol* (2008) **9**(6):439-446.

149. Doerner A, Rhiel L, Zielonka S, Kolmar H: **Therapeutic antibody engineering by high efficiency cell screening.** *FEBS letters* (2014) **588**(2):278-287.

150. Boder ET, Wittrup KD: **Yeast surface display for screening combinatorial polypeptide libraries.** *Nature biotechnology* (1997) **15**(6):553-557.

151. Sato N, Matsumoto T, Ueda M, Tanaka A, Fukuda H, Kondo A: **Long anchor using flo1 protein enhances reactivity of cell surface-displayed glucoamylase to polymer substrates.** *Appl Microbiol Biotechnol* (2002) **60**(4):469-474.

152. Ueda M, Tanaka A: **Genetic immobilization of proteins on the yeast cell surface.** *Biotechnol Adv* (2000) **18**(2):121-140.

153. Rakestraw JA, Aird D, Aha PM, Baynes BM, Lipovsek D: **Secretion-and-capture cell-surface display for selection of target-binding proteins.** *Protein engineering, design & selection : PEDS* (2011) **24**(6):525-530.

154. Shaheen HH, Prinz B, Chen MT, Pavoor T, Lin S, Houston-Cummings NR, Moore R, Stadheim TA, Zha D: **A dual-mode surface display system for the maturation and production of monoclonal antibodies in glyco-engineered pichia pastoris.** *PloS one* (2013) **8**(7):e70190.

155. Shusta EV, Kieke MC, Parke E, Kranz DM, Wittrup KD: **Yeast polypeptide fusion surface display levels predict thermal stability and soluble secretion efficiency.** *Journal of molecular biology* (1999) **292**(5):949-956.

156. Chao G, Lau WL, Hackel BJ, Sazinsky SL, Lippow SM, Wittrup KD: **Isolating and engineering human antibodies using yeast surface display.** *Nature protocols* (2006) **1**(2):755-768.

157. Bowley DR, Labrijn AF, Zwick MB, Burton DR: **Antigen selection from an HIV-1 immune antibody library displayed on yeast yields many novel antibodies compared to selection from the same library displayed on phage.** *Protein engineering, design & selection : PEDS* (2007) **20**(2):81-90.

158. Patriarca C, Macchi RM, Marschner AK, Mellstedt H: **Epithelial cell adhesion molecule expression (CD326) in cancer: A short review.** *Cancer treatment reviews* (2012) **38**(1):68-75.

159. Pavsic M, Guncar G, Djinovic-Carugo K, Lenarcic B: **Crystal structure and its bearing towards an understanding of key biological functions of EpCAM.** *Nature communications* (2014) **5**(4764.

160. Ladwein M, Pape UF, Schmidt DS, Schnolzer M, Fiedler S, Langbein L, Franke WW, Moldenhauer G, Zoller M: **The cell-cell adhesion molecule EpCAM interacts directly with the tight junction protein claudin-7.** *Exp Cell Res* (2005) **309**(2):345-357.

161. Lei Z, Maeda T, Tamura A, Nakamura T, Yamazaki Y, Shiratori H, Yashiro K, Tsukita S, Hamada H: **EpCAM contributes to formation of functional tight junction in the intestinal epithelium by recruiting claudin proteins.** *Dev Biol* (2012) **371**(2):136-145.

162. Winter MJ, Nagelkerken B, Mertens AE, Rees-Bakker HA, Briaire-de Bruijn IH, Litvinov SV: **Expression of Ep-CAM shifts the state of cadherin-mediated adhesions from strong to weak.** *Exp Cell Res* (2003) **285**(1):50-58.

163. Imrich S, Hachmeister M, Gires O: **EpCAM and its potential role in tumor-initiating cells.** *Cell adhesion & migration* (2012) **6**(1):30-38.

164. Ni J, Cozzi PJ, Duan W, Shigdar S, Graham PH, John KH, Li Y: **Role of the EpCAM (CD326) in prostate cancer metastasis and progression.** *Cancer metastasis reviews* (2012) **31**(3-4):779-791.

165. Kolibaba KS, Druker BJ: **Protein tyrosine kinases and cancer.** *Biochim Biophys Acta* (1997) **1333**(3):F217-248.

166. Gschwind A, Fischer OM, Ullrich A: **The discovery of receptor tyrosine kinases: Targets for cancer therapy.** *Nature reviews Cancer* (2004) **4**(5):361-370.

167. Tandon M, Vemula SV, Mittal SK: **Emerging strategies for EphA2 receptor targeting for cancer therapeutics.** *Expert opinion on therapeutic targets* (2011) **15**(1):31-51.

168. Pasquale EB: **Eph-ephrin bidirectional signaling in physiology and disease.** *Cell* (2008) **133**(1):38-52.

169. Kinch MS, Carles-Kinch K: **Overexpression and functional alterations of the EphA2 tyrosine kinase in cancer.** *Clinical & experimental metastasis* (2003) **20**(1):59-68.

170. Walker-Daniels J, Hess AR, Hendrix MJ, Kinch MS: **Differential regulation of EphA2 in normal and malignant cells.** *The American journal of pathology* (2003) **162**(4):1037-1042.

171. Zarogoulidis P, Katsikogianni F, Tsiouda T, Sakkas A, Katsikogiannis N, Zarogoulidis K: **Interleukin-8 and interleukin-17 for cancer.** *Cancer Invest* (2014) **32**(5):197-205.

172. Gales D, Clark C, Manne U, Samuel T: **The chemokine CXCL8 in carcinogenesis and drug response.** *ISRN oncology* (2013) **2013**(859154.

173. Verbeke H, Geboes K, Van Damme J, Struyf S: **The role of CXC chemokines in the transition of chronic inflammation to esophageal and gastric cancer.** *Biochim Biophys Acta* (2012) **1825**(1):117-129.

174. Kotyza J: **Interleukin-8 (CXCL8) in tumor associated non-vascular extracellular fluids: Its diagnostic and prognostic values. A review.** *The International journal of biological markers* (2012) **27**(3):169-178.

175. Singh JK, Farnie G, Bundred NJ, Simoes BM, Shergill A, Landberg G, Howell SJ, Clarke RB: **Targeting CXCR1/2 significantly reduces breast cancer stem cell activity and increases the efficacy of inhibiting HER2 via HER2-dependent and -independent mechanisms.** *Clin Cancer Res* (2013) **19**(3):643-656.

176. Huppa JB, Davis MM: **The interdisciplinary science of T-cell recognition.** *Advances in immunology* (2013) **119**(1-50.

177. Smith-Garvin JE, Koretzky GA, Jordan MS: **T cell activation.** *Annual review of immunology* (2009) **27**(591-619.

178. Fournier P, Schirrmacher V: **Bispecific antibodies and trispecific immunocytokines for targeting the immune system against cancer: Preparing for the future.** *BioDrugs : clinical immunotherapeutics, biopharmaceuticals and gene therapy* (2013) **27**(1):35-53.

179. Nagorsen D, Kufer P, Baeuerle PA, Bargou R: **Blinatumomab: A historical perspective.** *Pharmacol Ther* (2012) **136**(3):334-342.

180. Haas C, Krinner E, Brischwein K, Hoffmann P, Lutterbuse R, Schlereth B, Kufer P, Baeuerle PA: **Mode of cytotoxic action of t cell-engaging bite antibody MT110.** *Immunobiology* (2009) **214**(6):441-453.

181. Benatuil L, Perez JM, Belk J, Hsieh CM: **An improved yeast transformation method for the generation of very large human antibody libraries.** *Protein engineering, design & selection : PEDS* (2010) **23**(4):155-159.

182. VanAntwerp JJ, Wittrup KD: **Fine affinity discrimination by yeast surface display and flow cytometry.** *Biotechnology progress* (2000) **16**(1):31-37.

183. Wentzel A, Christmann A, Kratzner R, Kolmar H: **Sequence requirements of the gpng beta-turn of the *Ecballium elaterium* trypsin inhibitor II explored by combinatorial library screening.** *The Journal of biological chemistry* (1999) **274**(30):21037-21043.

184. Zielonka S: **Shark attack - Immunisierung von Haien zur Gewinnung von Eindomänen-antikörpermolekülen für den Einsatz in Diagnostik und Therapie.** *Fachbereich Chemie, Technische Universität Darmstadt* (2011) **(Masterthesis Biomolecular Engineering)**

185. Wu TT, Johnson G, Kabat EA: **Length distribution of CDRH3 in antibodies.** *Proteins* (1993) **16**(1):1-7.

186. Zemlin M, Klinger M, Link J, Zemlin C, Bauer K, Engler JA, Schroeder HW, Jr., Kirkham PM: **Expressed murine and human CDR-H3 intervals of equal length exhibit distinct repertoires that differ in their amino acid composition and predicted range of structures.** *Journal of molecular biology* (2003) **334**(4):733-749.

187. Zielonka S, Weber N, Becker S, Doerner A, Christmann A, Christmann C, Uth C, Fritz J, Schafer E, Steinmann B, Empting M *et al*: **Shark attack: High affinity binding proteins derived from shark vNAR domains by stepwise in vitro affinity maturation.** *Journal of biotechnology* (2014) **191**(236-245.

188. Grzeschik J: **Isolation and affinity maturation of single-somain antibodies for applications in cell biology and protein crystallography.** *Fachbereich Chemie, Technische Universität Darmstadt* (2014) **Masterthesis (Biomolecular Engineering)**

189. Lipovsek D, Lippow SM, Hackel BJ, Gregson MW, Cheng P, Kapila A, Wittrup KD: **Evolution of an interloop disulfide bond in high-affinity antibody mimics based on fibronectin type III domain and selected by yeast surface display: Molecular convergence with single-domain camelid and shark antibodies.** *Journal of molecular biology* (2007) **368**(4):1024-1041.

190. Heideman DA, Snijders PJ, Craanen ME, Bloemena E, Meijer CJ, Meuwissen SG, van Beusechem VW, Pinedo HM, Curiel DT, Haisma HJ, Gerritsen WR: **Selective gene delivery toward gastric and esophageal adenocarcinoma cells via EpCAM-targeted adenoviral vectors.** *Cancer Gene Ther* (2001) **8**(5):342-351.

191. Prang N, Preithner S, Brischwein K, Goster P, Woppel A, Muller J, Steiger C, Peters M, Baeuerle PA, da Silva AJ: **Cellular and complement-dependent cytotoxicity of EpCAM-specific monoclonal antibody MT201 against breast cancer cell lines.** *Br J Cancer* (2005) **92**(2):342-349.

192. Sterzynska K, Kempisty B, Zawierucha P, Zabel M: **Analysis of the specificity and selectivity of anti-EpCAM antibodies in breast cancer cell lines.** *Folia histochemica et cytobiologica / Polish Academy of Sciences, Polish Histochemical and Cytochemical Society* (2012) **50**(4):534-541.

193. Kurosawa G, Akahori Y, Morita M, Sumitomo M, Sato N, Muramatsu C, Eguchi K, Matsuda K, Takasaki A, Tanaka M, Iba Y *et al*: **Comprehensive screening for antigens overexpressed on carcinomas via isolation of human mabs that may be therapeutic.** *Proceedings of the National Academy of Sciences of the United States of America* (2008) **105**(20):7287-7292.

194. Flieger D, Hoff AS, Sauerbruch T, Schmidt-Wolf IG: **Influence of cytokines, monoclonal antibodies and chemotherapeutic drugs on epithelial cell adhesion molecule (EpCAM) and lewisy antigen expression.** *Clin Exp Immunol* (2001) **123**(1):9-14.

195. Ponsel D, Neugebauer J, Ladetzki-Baehs K, Tissot K: **High affinity, developability and functional size: The holy grail of combinatorial antibody library generation.** *Molecules (Basel, Switzerland)* (2011) **16**(5):3675-3700.

196. Chan CE, Lim AP, MacAry PA, Hanson BJ: **The role of phage display in therapeutic antibody discovery.** *International immunology* (2014) **26**(12):649-657.

197. Hackel BJ, Wittrup KD: **The full amino acid repertoire is superior to serine/tyrosine for selection of high affinity immunoglobulin G binders from the fibronectin scaffold.** *Protein engineering, design & selection : PEDS* (2010) **23**(4):211-219.

198. Davies J, Riechmann L: **Antibody VH domains as small recognition units.** *Bio/technology (Nature Publishing Company)* (1995) **13**(5):475-479.

199. Davies J, Riechmann L: **Affinity improvement of single antibody VH domains: Residues in all three hypervariable regions affect antigen binding.** *Immunotechnology : an international journal of immunological engineering* (1996) **2**(3):169-179.

200. Ewert S, Honegger A, Pluckthun A: **Structure-based improvement of the biophysical properties of immunoglobulin VH domains with a generalizable approach.** *Biochemistry* (2003) **42**(6):1517-1528.

201. Ewert S, Huber T, Honegger A, Pluckthun A: **Biophysical properties of human antibody variable domains.** *Journal of molecular biology* (2003) **325**(3):531-553.

202. Feldhaus MJ, Siegel RW, Opresko LK, Coleman JR, Feldhaus JM, Yeung YA, Cochran JR, Heinzelman P, Colby D, Swers J, Graff C *et al*: **Flow-cytometric isolation of human antibodies from a nonimmune *Saccharomyces cerevisiae* surface display library.** *Nature biotechnology* (2003) **21**(2):163-170.

203. Glotzbach B, Reinwarth M, Weber N, Fabritz S, Tomaszowski M, Fittler H, Christmann A, Avrutina O, Kolmar H: **Combinatorial optimization of cystine-knot peptides towards high-affinity inhibitors of human matriptase-1.** *PloS one* (2013) **8**(10):e76956.

204. Stone JD, Chervin AS, Aggen DH, Kranz DM: **T cell receptor engineering.** *Methods in enzymology* (2012) **503**(189-222.

205. Walker LM, Bowley DR, Burton DR: **Efficient recovery of high-affinity antibodies from a single-chain Fab yeast display library.** *Journal of molecular biology* (2009) **389**(2):365-375.

206. Ellgaard L, Molinari M, Helenius A: **Setting the standards: Quality control in the secretory pathway.** *Science* (1999) **286**(5446):1882-1888.

207. Young NM, MacKenzie CR, Narang SA, Oomen RP, Baenziger JE: **Thermal stabilization of a single-chain Fv antibody fragment by introduction of a disulphide bond.** *FEBS letters* (1995) **377**(2):135-139.

208. Yasui H, Ito W, Kurosawa Y: **Effects of substitutions of amino acids on the thermal stability of the Fv fragments of antibodies.** *FEBS letters* (1994) **353**(2):143-146.

209. Worn A, Pluckthun A: **Stability engineering of antibody single-chain Fv fragments.** *Journal of molecular biology* (2001) **305**(5):989-1010.

210. Boder ET, Raeeszadeh-Sarmazdeh M, Price JV: **Engineering antibodies by yeast display.** *Archives of biochemistry and biophysics* (2012) **526**(2):99-106.

211. Gera N, Hussain M, Rao BM: **Protein selection using yeast surface display.** *Methods (San Diego, Calif)* (2013) **60**(1):15-26.

212. Bumbaca D, Wong A, Drake E, Reyes AE, 2nd, Lin BC, Stephan JP, Desnoyers L, Shen BQ, Dennis MS: **Highly specific off-target binding identified and eliminated during the humanization of an antibody against FGF receptor 4.** *mAbs* (2011) **3**(4):376-386.

213. Feyen O, Lueking A, Kowald A, Stephan C, Meyer HE, Gobel U, Niehues T: **Off-target activity of TNF-alpha inhibitors characterized by protein biochips.** *Anal Bioanal Chem* (2008) **391**(5):1713-1720.

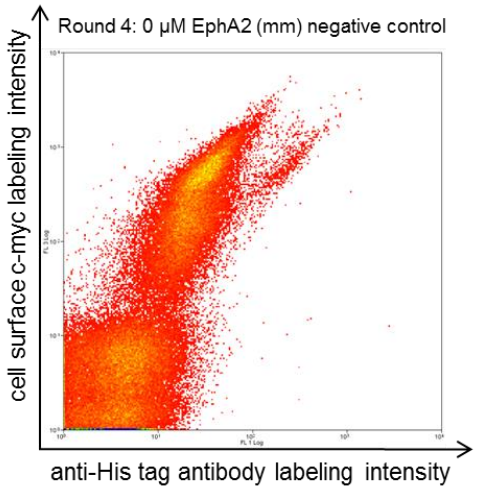
214. Wu H, Pfarr DS, Johnson S, Brewah YA, Woods RM, Patel NK, White WI, Young JF, Kiener PA: **Development of Motavizumab, an ultra-potent antibody for the prevention of respiratory syncytial virus infection in the upper and lower respiratory tract.** *Journal of molecular biology* (2007) **368**(3):652-665.

215. Halaby DM, Poupon A, Mornon J: **The immunoglobulin fold family: Sequence analysis and 3D structure comparisons.** *Protein engineering* (1999) **12**(7):563-571.

216. Xiao X, Feng Y, Vu BK, Ishima R, Dimitrov DS: **A large library based on a novel (CH2) scaffold: Identification of HIV-1 inhibitors.** *Biochem Biophys Res Commun* (2009) **387**(2):387-392.

8. APPENDIX

Appendix A. Negative control of screening round 4 against receptor tyrosine kinase EphA2



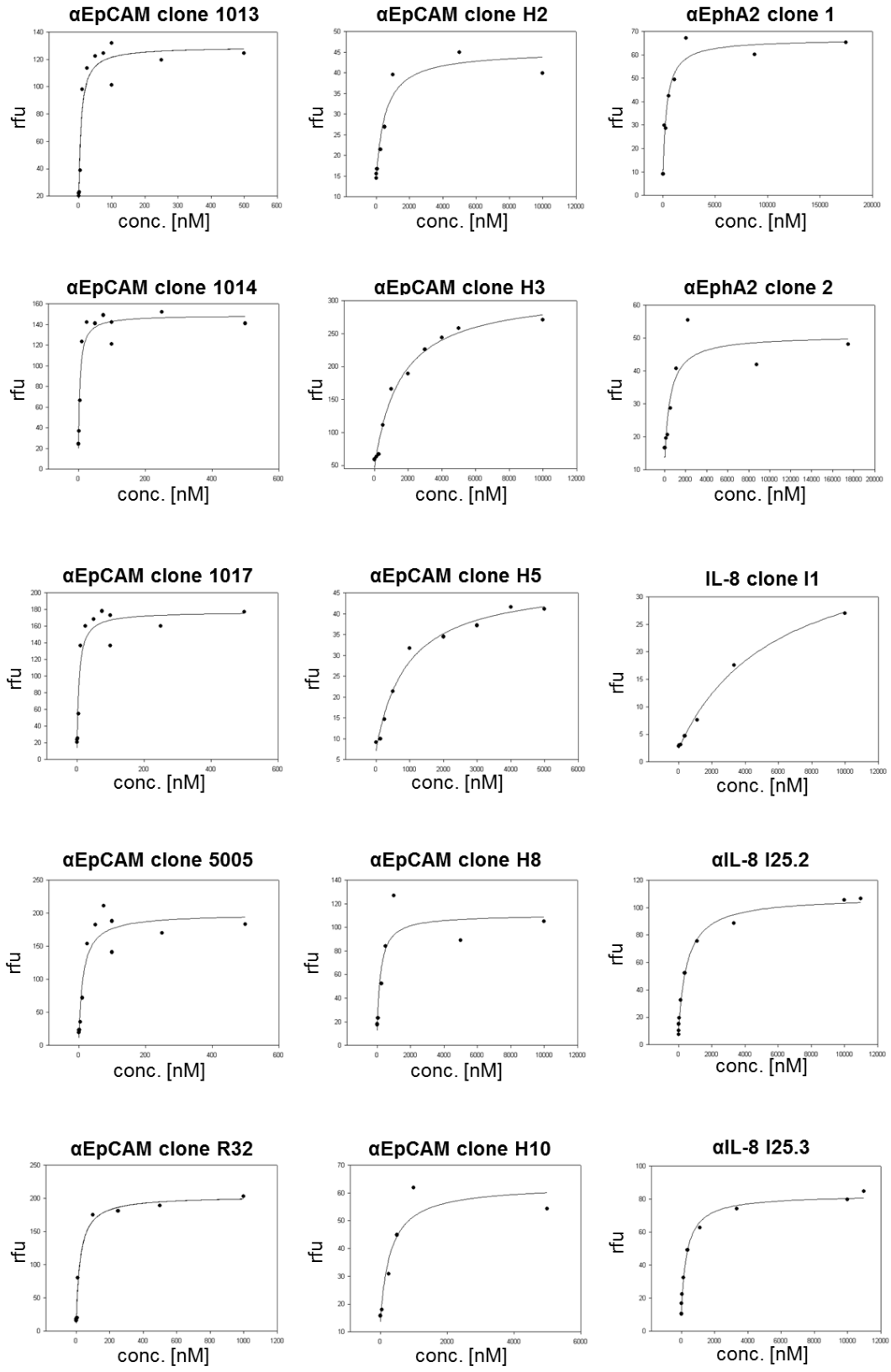
Appendix A. Negative control of Library screening round 4 against EphA2. Yeast cells were labeled for simultaneous detection of surface presentation and for detection of binding against Alexa Fluor 488 labeled anti-His-tag antibody.

Appendix B. Multiple sequence alignment of vNAR molecules of different types from different species

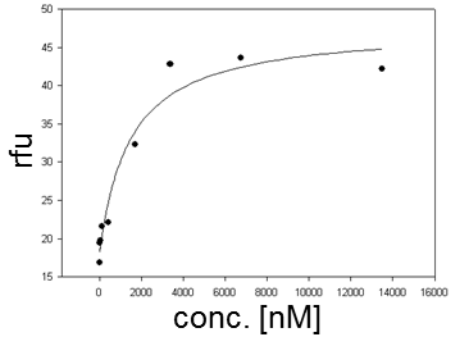
```
                                                    CDR1                HV2                 HV4
C. plagiosum Type IV  MAARLEQTPTTTTKEAGESLTINCVLKPEWTILGRTYWYFTKKGATKKARLSTGGRYSDTKNTASKSLSLRISDLRVEDS
C. plagiosum Type IV  AAARLEQTPTTTTKEAGESLTINCALKGSTYGLGSTYWYFTKKGATKKARLSTGGRYSDTKNTASKSFSLRISDLRVEDS
C. plagiosum Type II  MAQRVEQTPTTTTKEAGESLTINCVLKGSRCALGNTYWYFTKKGATKKARLSTRGRYSDTKNTASKSFSLRISDLRVEDS
G. cirratum Type II   --ARVDQTPRSVTKETGESLTINCVLRDKDCAESSASWYRKKSGSTNEESISKGGRYVETVNSGSKSFSLRINDLTVEDS
G. cirratum Type I    --ARVDQTPRSVTKETGESLTINCVLRDASYALGSTCWYRKKSGEGNEESISKGGRYVETVNSGSKSFSLRINDLTVEDG
S. acanthias Type IV  --ARVDQTPRTATKETGESLTINCVLTDTSYPLYSTYWYRKNPGSSNKEQISISGRYVESVNKGTKSFSLRIKDLTVADS
O. maculatus Type II  --ARVDQTPRIATKETGESLTINCVLRDTACALDSTNWYRTKLGSTKEQTISIGGRYSETVDEGSNSASLTIRDLRVEDS
T. scyllium Type II   --AHVDQTPRVATKETGESLTINCVLRGVSCGLYATRWFRQNPGSTSWERITIGGRYVESVNKGSKSFSLQIKDLTVEDS
O. maculatus Type III --ARVDQTPKTITKETGESLTINCVLSDTSCAWDSTYWYRKKLDSTNEESTSKGGRYVETVNSESTSFSLRINDLTVEDS
                        :::***   ***:********.*        : *:  : .  .    :  *** :: :  :.* ** * ** * *.

C. plagiosum Type IV  GTYHCEALIYSDM-----GM-I-MW-KIEGGGTTVTVK
C. plagiosum Type IV  GTYHCEAKAGHWF-----NL-L-HF-NIEGGGTTVTVK
C. plagiosum Type II  GTYHCEAYTAGMT-----NC-W-NS-YIEGGGTPLTVK
G. cirratum Type II   GTYRCKVPSRYSY--DCVRFEL-ID-DVYGDGTAVTV-
G. cirratum Type I    GTYRCGLGVAGGY--CDYALCS-SRYAECGDGTAVTVN
S. acanthias Type IV  ATYICRAMGT------------DSW-TGDGAGTVLTVN
O. maculatus Type II  GTYKCKAYRRCAF-----NT-G-VG-YKEGAGTVLTVK
T. scyllium Type II   VTFYCKAQENTEEYYVGDRRCSRSN-YYDGTGTVMTVN
O. maculatus Type III GTYRCRAYLYCGA-----EL-D-SF-DEYGGGTIVTVN
                       *: *                 * ** :**
                                  CDR3
```

Appendix B. Multiple sequence alignment (CLUSTAL W, v1.83) of different vNAR protein sequences indicating highly conserved Gly-Arg-Tyr motif at the end of HV2. Sequences of vNAR molecules were randomly chosen. CDR1, HV4 and CDR3 are shaded red, respectively. HV2 shaded blue and conserved motif shaded yellow.

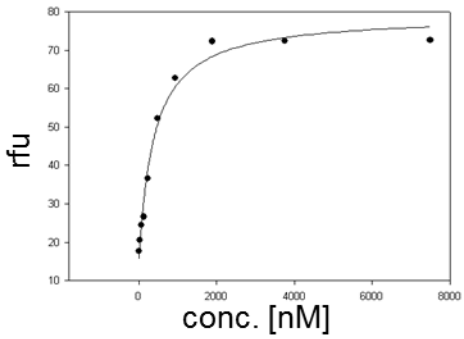
Appendix C. Affinities (K_D) determined by yeast surface display



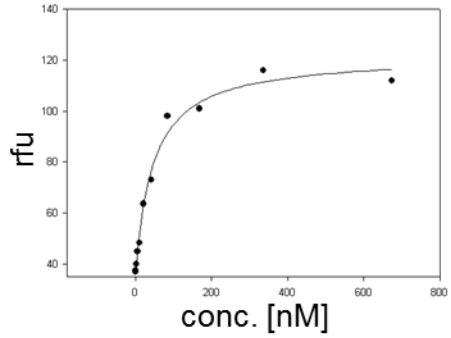
HV2-rand clone B1 (EpCAM)



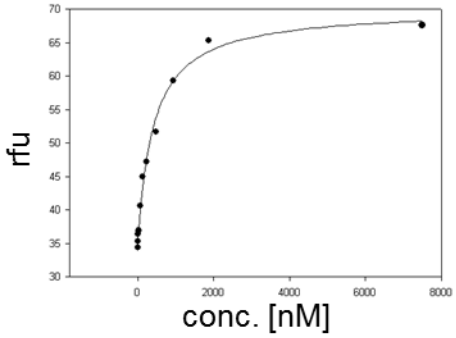
HV2-rand clone B1 (CD3ε)



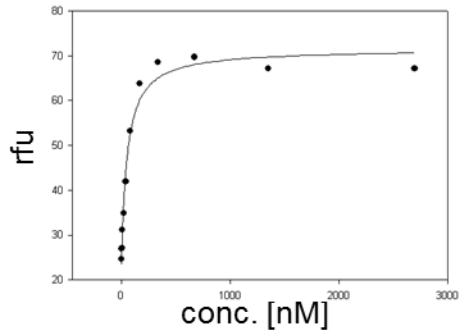
HV2-rand clone B1mut (EpCAM)



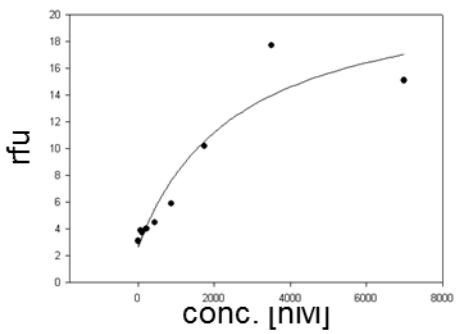
HV2-rand clone B1mut (CD3ε)



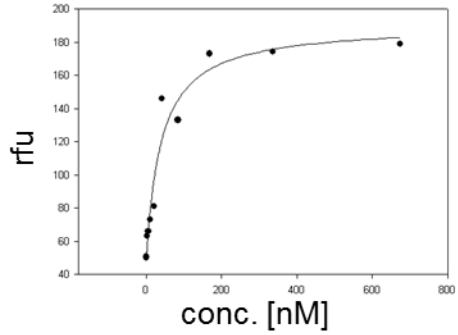
HV2-rand clone F1 (EpCAM)



HV2-rand clone F1 (CD3ε)



5005 (EpCAM)



Appendix D. Abbreviations

ADCC	Antibody-dependent cellular cytotoxicity
AMA1	Malarial apical membrane antigen-1
APC	Allophycocyanin
BCR	B-cell receptor
BiP	Ig-binding protein
BSA	Bovine serum albumin
bsAb	Bi-specific antibody
CD3ε	Cluster of differentiation 3ε
CDC	complement-dependent cytotoxicity
CDR	Complementarity determining region
CH2, CH3	Constant domain 2,3 of the heavy chain
CXCL8	Interleukin-8
DNL	Dock and Lock
Dscam molecule	Down syndrome cell adhesion molecule
EpCAM	Epithelial cell adhesion molecule
EphA2	Receptor tyrosine kinase EphA2
FACS	fluorescence-activated cell sorting
Fcγ	Fc-part of IgG1
FcRn	Neonatal Fc receptor
FDA	US Food and Drug Administration
FITC	Fluorescein isothiocyanate
FREP	Fibrinogen-related protein
Fw	Framework region
GFC	Gel filtration chromatography
GFP	Green fluorescent protein
HBeAg	Hepatitis B e antigen
HBV	Hepatitis B virus
HCAb	Heavy-chain only antibody
HSA	Human serum albumin
HTRA1	Human serine protease HTRA1
HV	Hypervariable loop
IgNAR	Immunoglobulin New Antigen Receptor
IgNAR V domain	IgNAR variable domain
IL-8	Interleukin-8
IMAC	Immobilized Metal Ion Affinity Chromatography

K_D	Equilibrium dissociation constant
k_{off}	Dissociation rate
LRR	Leucine-rich repeat
mAbs	Monoclonal antibodies
MalE	Maltose binding protein
MHC	Major Histocompatibility Complex
mm	murine
MWCO	Molecular weight cut off
PAGE	Polyacrylamide gel electrophoresis
PAMP	Pathogen associated molecular pattern
PCR	Polymerase chain reaction
PE	Phycoerythrin
PRR	Pathogen recognition receptor
rh	Recombinant human
scFv	Single-chain variable fragment
SDS	sodium dodecyl sulfate
SOE	Splicing by overlap extension
TCR	T-cell receptor
TEV	Tabacco etch virus
Tm	Melting temperature
TNF	Tumor necrosis factor
VH	Variable domain of the heavy chain
VHH	Variable domain of a camelid heavy-chain only antibody
VHSV	Viral hemorrhagic septicaemia virus
VL	Variable domain of the light chain
VLR	Variable lymphocyte receptors
vNAR	IgNAR variable domain
YSD	Yeast surface display

Appendix E. List of figures

Figure 1. Structural features of IgG antibody 6

Figure 2. Structural features of shark IgNAR antibody formats. 10

Figure 3. Comparison of VH and vNAR binding domains 12

Figure 4. Different types of IgNAR V domains 13

Figure 5. Translocon arrangements of immunoglobulin genes in higher vertebrates and cluster configuration of IgNAR genes of cartilaginous fish 15

Figure 6. Examples of CDR3 variability in vNAR domains depicted in ribbon representation. 17

Figure 7. Simplified illustration of yeast surface display according to Boder and Wittrup 21

Figure 8. Structure of a vNAR binding domain 25

Figure 9. Vector map for pCT plasmid 27

Figure 10. Vector map for pMX-EETI plasmid. 28

Figure 11. Vector map for pMX-vNAR plasmid 29

Figure 12. Vector map for pExpress-Fc plasmid 30

Figure 13. Analysis of the natural vNAR repertoire of the bamboo shark 51

Figure 14. Library construction and analysis 53

Figure 15. Distribution of amino acids in CDR3 of the unselected library 54

Figure 16. Library screening against EpCAM 55

Figure 17. Sequences of EpCAM-binding vNAR molecules selected from initial library screen. 55

Figure 18. Library screening against EphA2 57

Figure 19. Sequences of EphA2-binding vNAR molecules selected from initial library screen 58

Figure 20. Fourth round of library screening against IL-8. 59

Figure 21. Sequences of IL-8-binding vNAR molecules selected from initial library screen 59

Figure 22. Screening of an α-EpCAM sublibrary after randomization of CDR1 61

Figure 23. Sequences of selected α-EpCAM-vNARs after CDR1 randomization 61

Figure 24. Screening of an α-IL-8 sublibrary after randomization of CDR1 62

Figure 25. Sequences of selected α-IL-8-vNARs after CDR1 randomization 62

Figure 26. Scheme of of recombinant vNAR domain production as MalE-fusion protein 64

Figure 27. Elution profile of selected and recombinant produced MalE-vNAR fusion-protein on a His Trap 65

Figure 28. Purification of recombinant vNAR proteins *via* size exclusion chromatography 66

Figure 29. Elution profile of α-EpCAM_vNAR 5005-Fc fusion protein on a HiTrap Protein A HP 1 ml column 68

Figure 30. Binding curves of α-EpCAM vNARs used to determine K_D values 69
Figure 31. Thermal shift assay for T_m calculation of α-EpCAM vNARs 71
Figure 32. Specificity of α-EpCAM vNAR H3, H5 and 5005 72
Figure 33. Specificity of α-EpCAM vNAR 5005 Fc fusion protein validated by tumor cell staining assays 74
Figure 34. Depiction of the rationale for the generation of a new antigen-binding site into the vNAR scaffold 75
Figure 35. Schematic representation of PCR-based randomization of HV2 76
Figure 36. Analysis of HV2-randomized library based on EpCAM-specific vNAR 5005 77
Figure 37. Screening of a HV2-randomized EpCAM-specific library based on vNAR 5005 against CD3ε 78
Figure 38. Single clone analysis of potentially bi-specific vNARs after randomization of HV2 79
Figure 39. Sequences of potentially bispecific EpCAM-binding and CD3ε-binding vNARs 80
Figure 40. Single clone analysis of mono-specific, parental clone 5005 and HV2-diversified clone B1 81
Figure 41. Screening of a HV2-randomized EpCAM-specific library for bi-specificity against human Fcγ 82
Figure 42. Sequences of bispecific EpCAM-binding and Fcγ-binding vNAR F1 83
Figure 43. Single clone analysis of EpCAM- and glycosylated Fcγ-binding bi-specific clone F1 83

Appendix F. List of tables

Table 1. vNAR domains selected against disease related targets 18
Table 2. Equilibrium dissociation constants determined by yeast surface display 56
Table 3. Equilibrium dissociation constants determined by yeast surface display after affinity maturation 63
Table 4. Binding kinetics and Tm of soluble vNAR variants selected against EpCAM 70
Table 5. Different types of IgNAR V domains 79

Appendix G. *Curriculum vitae*

M.Sc.

Stefan Zielonka

Am Elfengrund 65
64297 Darmstadt
Germany
E-Mail: Zielonka@biochemie-TUD.de

Personal Data

Birth:	July 4th, 1986 in Viernheim (Germany)
Nationality:	German
Marital Status:	Unmarried

Postgraduate

January 2012 – March 2015	**Doctorate** at Technische Universität Darmstadt Clemens-Schöpf Institut für Organische Chemie und Biochemie, TU Darmstadt – Biochemistry Prof. Dr. Harald Kolmar

Graduate Degree

2010 - 2011	**Master of Science in Biomolecular Engineering**, Technische Universität Darmstadt Thesis: *Shark attack - Immunisierung von Haien zur Gewinnung von Eindomänen-Antikörpermolekülen für den Einsatz in Diagnostik und Therapie - Immunization of sharks for the generation of single domain antibody molecules* Prof. Dr. Harald Kolmar and Prof. Dr. Siegfried Neumann Final grade: 1,08

2007 - 2010	**Bachelor of Science in Biotechnology**, Hochschule Darmstadt 2007 - 2010, 6 Semester Grade point average: **1,2** (**With Honor**) Thesis: *Über die evolutive Optimierung von Antikörpervarianten in der Hefe Saccharomyces cerevisiae – AID mediated affinity maturation of antibody fragments in S. cerevisiae* Prof. Dr. Harald Kolmar and Prof. Dr. Stefan Hüttenhain Final grade: 1,2

School

1997 - 2006	**High school, Gymnasium** Starkenburg Gymnasium Heppenheim Final grade: 2,3
1993 - 1997	**Grundschule**, Grundschule Wingertsbergschule Lorsch

M.Sc. Stefan Zielonka

9. AFFIRMATIONS

Stefan Zielonka
Am Elfengrund 65
64297 Darmstadt

_____

Eidesstattliche Erklärung

Ich erkläre hiermit, dass ich meine Dissertation selbstständig und nur mit den angegebenen Hilfsmitteln angefertigt habe.

_____

Stefan Zielonka
Am Elfengrund 65
64297 Darmstadt

_____

Erklärung

Ich erkläre hiermit, noch keinen Promotionsversuch unternommen zu haben.

_____

Danksagung

Folgenden Personen möchte ich meinen aufrichtigen Dank aussprechen:

Meinem Doktorvater, **Prof. Dr. Harald Kolmar**, möchte ich für die massive Unterstützung meiner Person sowie für viele anregende Gespräche bedanken. Während meiner Dissertation und schon zu Bachelor- und Masterzeiten durfte ich durchweg interessante Themen bearbeiten und Harald ließ mir dabei immer den nötigen Freiraum, eigene Ideen in Projekten zu verwirklichen. Zudem durfte ich im Rahmen der Dissertation Konferenzen besuchen sowie bei einer Vorlesung als Vortragender aktiv teilnehmen. Vielen Dank dafür!

Prof. Dr. Siegfried Neumann möchte ich vielmals für die erhebliche wissenschaftliche Förderung meiner Person bedanken. Seine Gutachten haben mir so manche Tür geöffnet und in gemeinsamen Diskussionen durfte ich an seinem schier unerschöpflichen Wissensschatz teilhaben. Darüber hinaus danke ich Ihm für viele konkrete Hinweise, die mein Dissertationsvorhaben aktiv vorangebracht haben. Außerdem danke ich ihm für die Übernahme des ersten Fachprüfers der Disputation.

Prof. Dr. Stefan Dübel und **Prof. Dr. Thomas Worzfeld** danke ich für die Übernahme des Korreferenten bzw. zweiten Fachprüfers meiner Disputation.

Prof. Dr. Stefan Hüttenhain danke ich für viele anregende und freundschaftliche Gespräche. Darüber hinaus hat er den Grundstein für meine wissenschaftliche Entwicklung gelegt indem er mich enorm während der Bachelorzeit unterstützt hat. Dadurch wurde bei mir das Interesse an der Wissenschaft erst wirklich geweckt. Weiterhin möchte ich mich für das gemeinsame Kaffee-Trinken sowie die regelmäßigen Alumni-Treffen und Dreikönigstags-Veranstaltungen bedanken, die immer einen netten Rahmen bieten, ehemalige Kommilitonen wiederzusehen.

Stefan Becker danke ich für das *Mentoring* während des Promotionsstipendiums der Merck'schen Gesellschaft für Wissenschaft und Kunst sowie darüber hinaus. Durch seine konstruktiven Vorschläge sowie ebenfalls durch Unterstützungen in Form von Sachmitteln wurden die Projekte während der Dissertation sehr bereichert. **Achim Doerner** danke ich ebenfalls für konstruktive Vorschläge bezüglich meiner Dissertation sowie für das freundschaftliche Verhältnis. Dazu gehören einige gemeinsame Konzerterfahrungen, unzählige feucht-fröhliche Kochabende, gemeinsame Lauftreffs und noch so einiges mehr! **Laura Rhiel** danke ich für die Kooperation und das gemeinsame Schreiben des *Reviews*.

Björn Steinmann und **Martin Empting** danke ich für die Freundschaft, die sich während der gemeinsamen Zeit im Arbeitskreis entwickelt hat. Gerade die Anfangszeit für mich im Arbeitskreis wäre ohne euch kaum vorstellbar gewesen. Rückblickend denke ich hierbei an unzählige gemeinsame Ereignisse wie beispielsweise Konzerte, das betreute Trinken, feuchte Tabakerlebnisse, Schweigen-der-Lämmer-Gedenkmomente während Vorträgen sowie speziell mit Genosse Martin Empting an die durchweg intensiven politischen Debatten. Darüber hinaus möchte ich beiden für die Anleitung, die

Kooperation, für aktive Beiträge, Ratschläge und konstruktive Diskussionen in den verschiedenen Projekten meiner Dissertation danken. Hier gilt nochmals ein großes Dank an Martin Empting, der prinzipiell an allen Projekten aktiv beteiligt war und auch konzeptionell Projekte mitplante. In diesem Zusammenhang möchte ich ebenfalls **Pia Ockelmann** für die Kooperation danken, sowie **Prof. Dr. Lierz**.

Janine Becker danke ich für die Unterstützung und für das „Anlernen" gerade in der Anfangszeit als Bachelorstudent. Janine hat mir gezeigt, dass *E. coli*s auf dem Kopf wachsen. Weiterhin möchte ich mich für generelle Hilfen bezüglich Klonierungen, Sucrose-Gradienten und vor allem für die aktive Durchführung der *Deep Sequencing* Projekte bedanken, sowie für einige Kaltschalen, die wir uns gemeinsam bei unterschiedlichen *Social Events* und beim *Pirate Sattelite Festival 2012* genehmigen durften. **Barbara Diestelmann** danke ich für die Hilfe in vielen organisatorischen Fragen, für das Bereitstellen von Laborutensilien sowie für die vielen anregenden Gespräche während Zigarettenpausen und für das gemeinsame Verarbeiten nach einer Dokumentation.

Christina Uth danke ich für die Freundschaft, die sich hier im Arbeitskreis entwickelt hat, für die gemeinsamen Projekte während der Dissertation, für die unzähligen Konzertabende, Koch-, Grill- und Trinkabende, Tischkickerspiele und für gemeinsame Konferenzbesuche, allem voran die *PEGS Europe* in Wien.

Thomas Hofmeyer aka *Musclemeyer* aka *Bud Sperkulegger* danke ich für die *Benchbruderschaft*. Die Zeit neben dir und Niklas war wirklich legendär. Ich danke Thomas dafür, dass ich ihn nach einigen Remis-Partien nun gänzlich in einem hart umkämpften Armdrückduell bezwungen habe. Weiterhin danke ich ihm für gemeinsame Wanderungen, für unzählige gemeinsame *Pivo,* für gemeinsame Erlebnisse im *Bruce Willis Actionschacht* sowie für ein ganz besonders unkonventionelles Fechtduell, das ich mein Leben lang nicht vergessen werde.

Niklas Weber aka *Niklafar* aka *Sicknick* danke für die gemeinsame Freundschaft und für die *Benchbruderschaft*. Ich danke dir für die Kooperation im Haiprojekt und die gemeinsamen Konferenzbesuche, hier sei nur kurz Berlin zu erwähnen, wobei alle Interna verschwiegen werden müssen! Ich danke dir für das Erlebnis im *Bruce Willis Actionschacht*, für gemeinsame Wanderungen, Konzerte, Touren nach Dublin, Hamburg, Prag und bald Lissabon, für unzählige Kochabende und dafür, dass du meinen musikalischen Horizont erweitert hast. Ich danke dir für gemeinsame Büdchentouren und so vieles mehr, das mir gerade nicht einfällt!

Ich danke **Simon Krah** und **Christian Schröter** für die gemeinsame Zeit in Lissabon, für gemeinsame Biere, für das gemeinsame Schreiben des *Reviews* sowie für diverse Unterstützung während meiner Dissertation. *Die allerbesten Jungs…*

Bernhard Valldorf danke ich für gemeinsame Biertouren durch Frankfurt, für das Besuchen der Vorträge im Georg-Speyer Haus sowie für die gemeinsame Zeit in Lissabon. Was wäre nur Lissabon

ohne dich für ein Höllentrip geworden. Mit dir hat es riesig Spaß gemacht, danke für das Babysitten während des Flugs und unmittelbar vor meinem Vortrag, danke für das Erkunden der Stadt, dafür dass du mich an deinem Bier hast riechen lassen und umgekehrt und für den Besuch in der pinken Straße. Weiterhin danke ich dir für die vielen Gespräche bezüglich arbeitsrelevanter Themen. Sagres, slat!

Stephan Dickgießer danke ich für die Hilfe und das „Anlernen" in der Zellkultur. Weiterhin danke ich ihm dafür, dass er meinen musikalischen Horizont nicht erweitert hat. Durch ihn habe ich gelernt, dass es wirklich objektive Kriterien für gute Musik gibt und dass er diese mit Füßen tritt. Gönn dir, Bruder! **Heiko Fittler** danke ich dafür, dass er der *Superchamp* ist sowie für das gemeinsame Tischfußball spielen.

Julius Grzeschick, **Doreen Könning**, **Sebastian Meister, Kerstin Brettschneider** und **Elena Schäfer** danke ich für die nette Zusammenarbeit während ihrer Master- bzw. Bachelorarbeit. Darüber hinaus danke ich Doreen und Julius für ihren Beitrag zum gemeinsamen Review sowie für arbeitsrelevante Diskussionen. Julius danke ich vor allem für Ivan Mladek, für unzählige gemeinsame Biere, für Diskussionen, gemeinsame Wanderungen, Prag, Papa Schlumpf und das Knusperhäuschen, das etwas andere *Sightseeing* in Wien, für den gemeinsamen Besuch der Vorträge in Frankfurt, sowie für das freundschaftliche Verhältnis!

Andreas Christmann danke ich für Vorschläge bezüglich der Dissertation, für die *Deep Sequencing* Projekte sowie für deren Auswertung.

Weiterhin danke ich folgenden Mitarbeitern und ehemaligen Doktoranden des AK Kolmars, ich hoffe man verzeiht mir, wenn ich die ein oder andere Person aufgrund der Vielzahl an zu dankenden Personen vergessen habe: **Bernhard Glotzbach, Caro Mai, Tim Heiseler, Alex Maass, Franzi Maas, Vanessa Siegmund, Sascha Knauer, Daichi Nasu, Sebastian Fabritz, Michael Reinwarth, Olga Avrutina**.

Ganz besonders möchte ich meiner Familie, sowie meiner zukünftigen Familie für den enormen Rückhalt während des Studiums und der Promotionszeit danken. Hierbei gilt mein Dank zuerst **meinem Vater**, der mich Zeit meines Lebens durchweg unterstützt hat. Der Dank hierfür ist mit Worten nicht zu umfassen! Danken möchte ich auch meinem Bruder **Matthias** sowie dessen Lebensgefährtin **Mareen** für unbeschreiblich viele Dinge, sowie ebenfalls für fachbezogene Gespräche. Weiterhin möchte ich **Monika Conrad**, **Helmut Conrad** sowie **Christoph Conrad** danken! Ganz besonders möchte ich meiner Freundin und Verlobten **Franziska Conrad** für so ziemlich alles und jeden gemeinsamen Moment danken: Für den besonderen Rückhalt und ihre Loyalität, für das Ertragen meiner Launen während des Studiums und der Promotion, für unendlich viele *schönste Tage* in meinem Leben, aber auch für das gemeinsame Überstehen schlechter Zeiten, in denen es ihr immer wieder gelang, meine Motivation wiederherzustellen. Im Voraus danke ich für alle Dinge, die uns in Zukunft noch erwarten werden!

Abschließend möchte ich nicht vergessen, meinen Freunden zu danken, die mich schon seit sehr langer Zeit begleiten. Allen voran möchte ich **Max Halbritter**, **Jojo Lamp** sowie **Robert Heinz** für unschätzbar viele und gleichermaßen wertvolle Erlebnisse und Erfahrungen danken. Zu Dank verpflichtet bin ich euch aus so vielen Gründen, um nur einige zu nennen: die vielen tiefsinnigen, aber auch stumpfsinnigen Gespräche, die unzähligen Biertouren in Frankfurt, Prag, Hamburg, Wien, Duisburg, Kassel, Köln, Dublin, Bremen, Los Angeles, San Diego, Las Vegas, Zürich, Paris, London und bald Lissabon, die gemeinsamen Wanderungen, die vielen Konzerte und Festivals, das Erweitern meines musikalischen Horizonts und vieles mehr. In diesem Zusammenhang möchte ich auch **Reiner Hessler** und **Daniel Kind** danken, die ich im Bachelorstudium an der Hochschule Darmstadt kennenlernen durfte und die mittlerweile zu wirklichen Freunden geworden sind. Ich danke euch nicht zuletzt für euren Besuch in den USA und damit verbunden für den Road Trip und das Steel Panther *Happening* in Las Vegas. Reiner danke ich darüber hinaus für Prag und zukünftig für Lissabon und damit verbunden für die vielen feucht-fröhlichen Momente sowie dafür, dass er sich als „Helikopter-Landeplatz“ geopfert hat!